sempre:
Society for Education, Music and Psychology Research

Article

Psychology of Music
40(4) 429–470

DOI: 10.1177/0305735610392103
pom.sagepub.com
SAGE

An empirical exploration of the zygonic model of expectation in music

Michael Thorpe
Roehampton University, UK

Adam Ockelford
Roehampton University, UK

Aleksandar Aksentijevic
Roehampton University, UK

Abstract

Aspects of the 'zygonic' model of expectation in music (Ockelford, 2006) were tested experimentally. Forty subjects were played a diatonic melody, starting with the initial note only, then the first two notes, and so on. Each time, subjects were asked to sing what they considered to be the most likely continuation. The results were compared with the outputs of three algorithms derived from the zygonic model, which took into account adjacency ('Z1'), adjacency and recency ('Z2'), and adjacency, recency, and between-group projections ('Z3'). Each algorithm modelled the perceptual responses with statistically distinct degrees of accuracy; Z3 was the most faithful to subjects' expectations. Given the empirical data, potential refinements to the quantification of the zygonic model were considered. Additionally, it was found that men and women exhibited different patterns of expectation in relation to the stimuli that were presented, paralleling recent neuropsychological data suggesting that the location of music-structural processing in the brain may differ by gender.

Keywords

adjacency, anticipation, musical structure, recency

Introduction

As David Huron so vividly describes, expectation pervades the human condition (2006, p. 3): 'A cook expects a broth to taste in a certain way. A pedestrian expects traffic to move when the light turns green. A poker player expects an opponent to bluff.' Indeed, the capacity to make

Corresponding author:
Michael Thorpe, Department of Psychology, Roehampton University, Whitelands College, Holybourne Avenue, London SW15 4JD, UK.
Email: thorpem@roehampton.ac.uk

judgements about the future based on past and present experience is apparently wired deep in our neural architecture (Kveraga, Ghuman, & Bar, 2007). Hence it is unsurprising that the importance of expectation has been reported by researchers working across the cognitive sciences: in visual perception, for instance (Engel, Fries, & Singer, 2001; Summerfield & Egner, 2009), attention (Hogarth, Dickinson, Austin, Brown, & Duka, 2008), memory (Whittlesea, Masson, & Hughes, 2005), language processing (Otten, Nieuwland, & van Berkum, 2007), and behaviour (Olson, Roese, & Zanna, 1996).

Expectation is implicit in much contemporary music theory too (Schmuckler, 1989, p. 111), often retracing the paths pioneered by Leonard Meyer (1956, 1967). Meyer's ideas, rooted in Gestalt perception and information theory, were extended by Eugene Narmour in his 'implication–realization' model (1977, 1990, 1992, 1996). This subsequently found some support in empirical studies (Schellenberg, 1996, 1997; Thompson & Stainton, 1996), which indicated that simplification leads to little diminution of its predictive power. Moreover, von Hippel and Huron's (2000) analysis of melodies from a variety of cultures showed that Narmour's key principle of 'registral return' could be explained as an artefact of constraints on range. Nonetheless, the broad thrust of his theory retains its relevance in certain areas of psychomusicological endeavour, in which other approaches to expectation also continue to play a prominent role (Bharucha, 1999; Jones, 1981, 1982, 1992; Margulis, 2003, 2005, 2007).

In 2006, Adam Ockelford drew a number of these strands of thinking together into the conceptual framework offered by 'zygonic' theory (Ockelford, 2005, 2006, 2008, 2009). This asserts that the cognition of musical structure stems from a sense of *derivation*, whereby musical elements are heard (typically nonconsciously) as existing in imitation of others. The relationships – hypothesized cognitive constructs – through which such derivation is held to occur are said to be 'zygonic' (from the Greek word for 'yoke', implying the union of two similar things). 'Zygons' constitute a type of 'interperspective relationship', through which perceived aspects of musical sounds are compared. Such relationships can be represented graphically as in Figure 1.[1]

Ockelford formulated a new model of expectation (2006, pp. 127ff), whereby anticipation in music is said to arise through the projection of zygonic relationships into the future, using what Husserl (1964) called 'protentions': the anticipation of what is to come, enauralized in the present. These relationships stem from one of two sources:

(a) 'current' structures, which form part of the hearing process in train at the time, are encoded in working memory, and operate either

  (i) *within* groups of notes or
  (ii) *between* them (**A** in Figure 2); and

(b) 'previous' structures (which formed part of past hearing processes, and therefore necessarily operate only *between* groups). These may be encoded 'schematically' (**B**) or 'veridically' (**C**) (see Bharucha, 1987, 1994).

Current 'within-group' structures can offer only a *general* indication of what is to come (**1**), since all musical events have a plurality of logical continuations. Conversely, 'between-group' expectation provides a *specific* indication of what is likely to happen next (**3**). Within current structures, prognostication may be prompted by features that are particularly salient, through their recency or frequent repetition. Schematic information derived from structures heard previously offers a *general* picture of what the future may hold (**2**), according to heuristics based on past trends and tendencies. In summary:

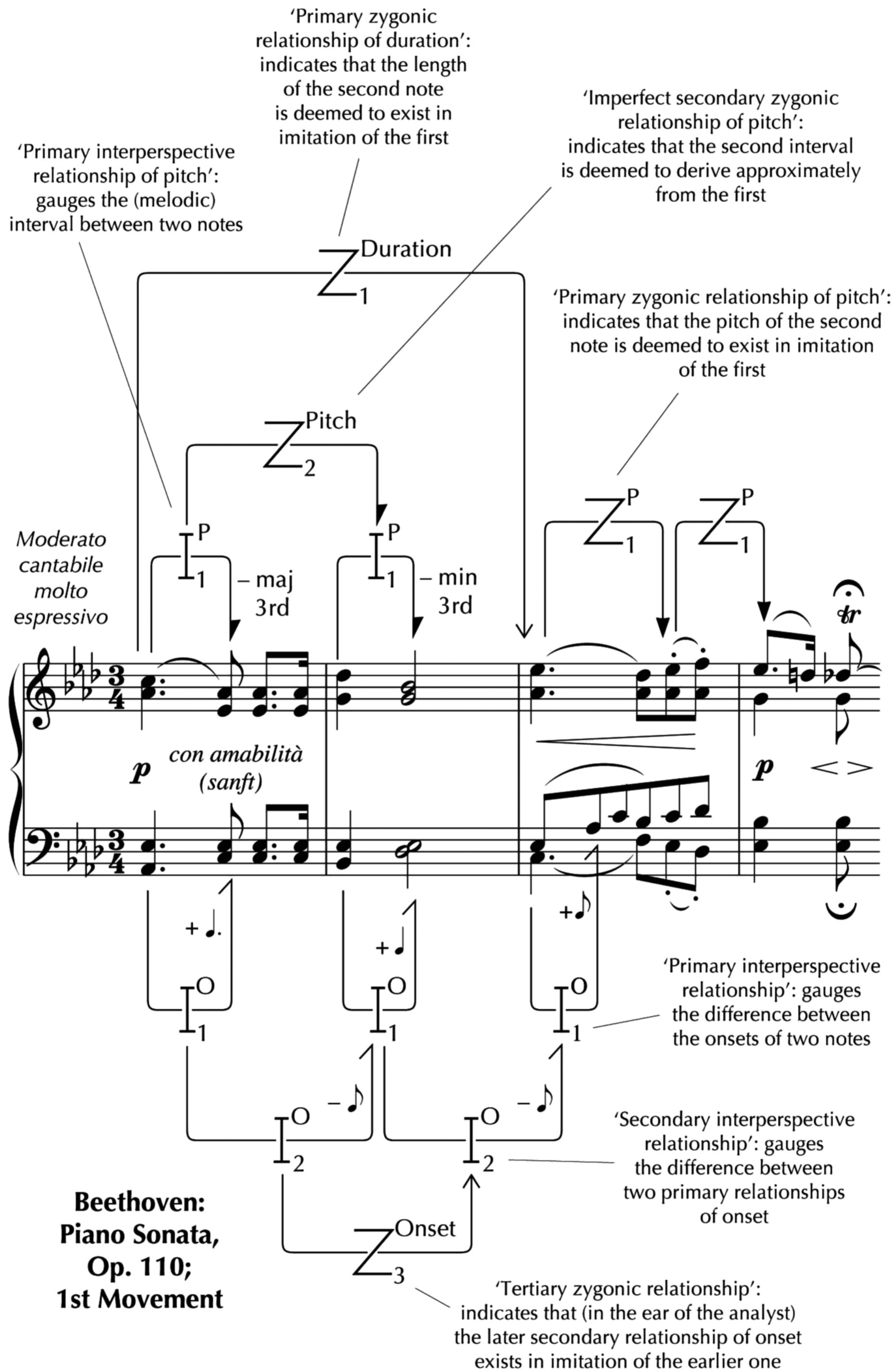


Figure 1. Examples of interperspective and zygonic relationships

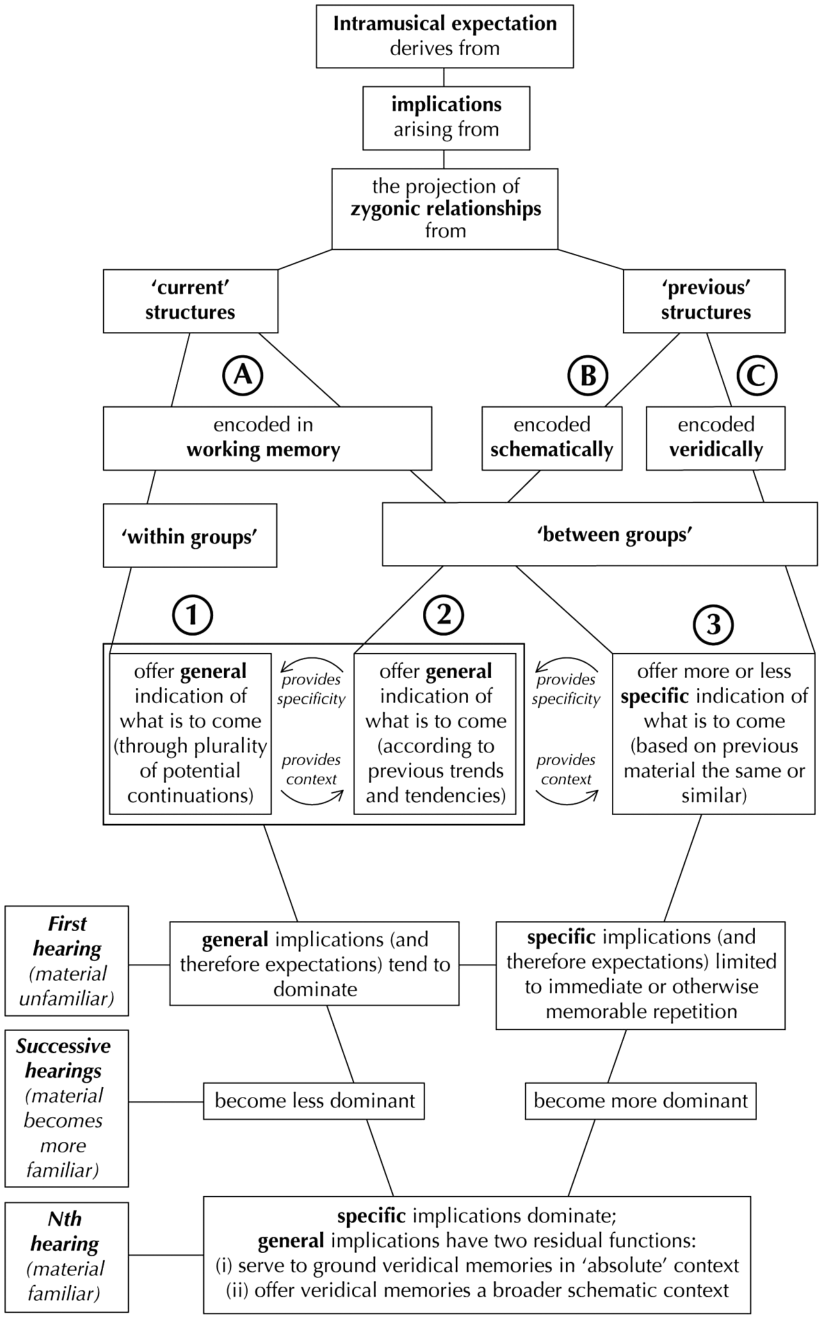


Figure 2. Zygonic model of expectation in music (after Ockelford, 2006, p.127)

A → **1**
B → **2**
A and/or **C** → **3**

1 and **2** interact, whereby **B** lends greater *specificity* to the implications arising from **A**, which provide a *local context* for the projections stemming from **B**. Consider, for example, the third movement of Rachmaninov's Symphony No. 2 (see Figure 3). Imagine a listener with musical experience *entirely outside* the Western tradition (scenario **A** alone), who has therefore not heard the work before. She listens to the first bar.

What pitch would she expect to occur next in the first violins? Projecting zygonic relationships forward suggests that *any* note framed by the semitonal universe with nodes at concert pitch would offer a logical option for continuation (see Figure 4).

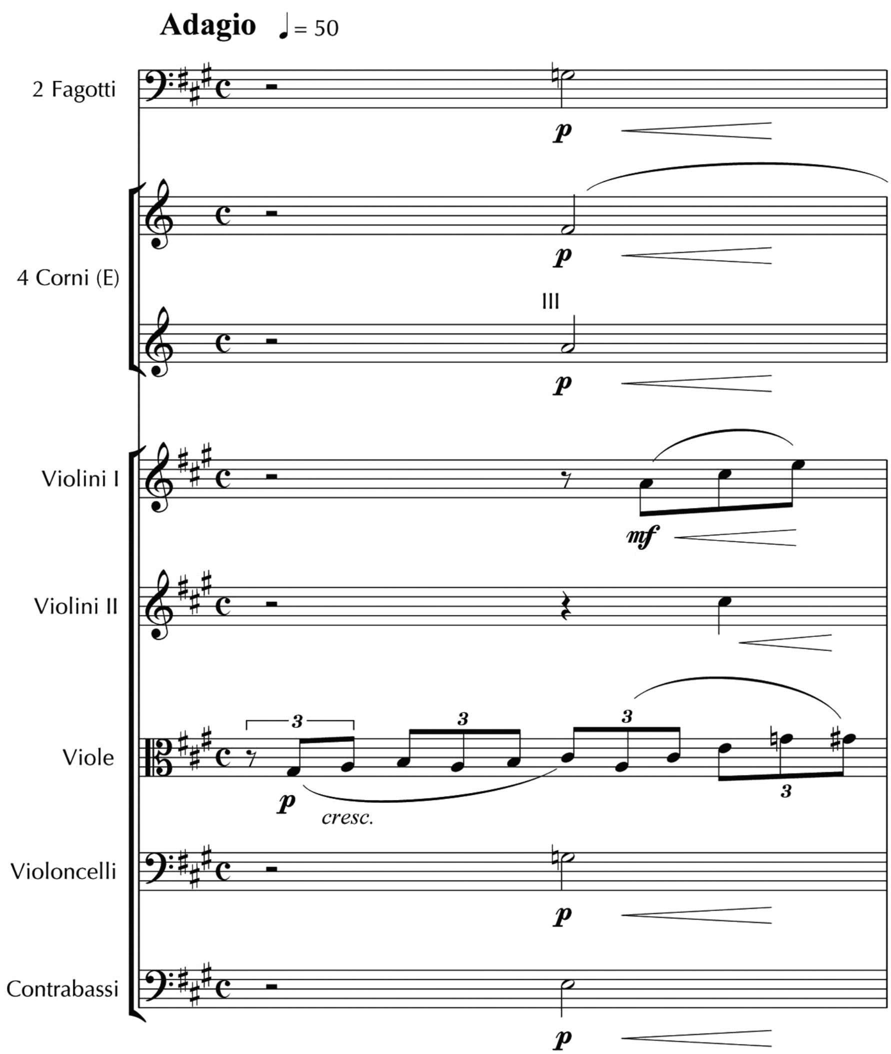


Figure 3. The opening bar of Rachmaninov's 2nd Symphony, 3rd Movement

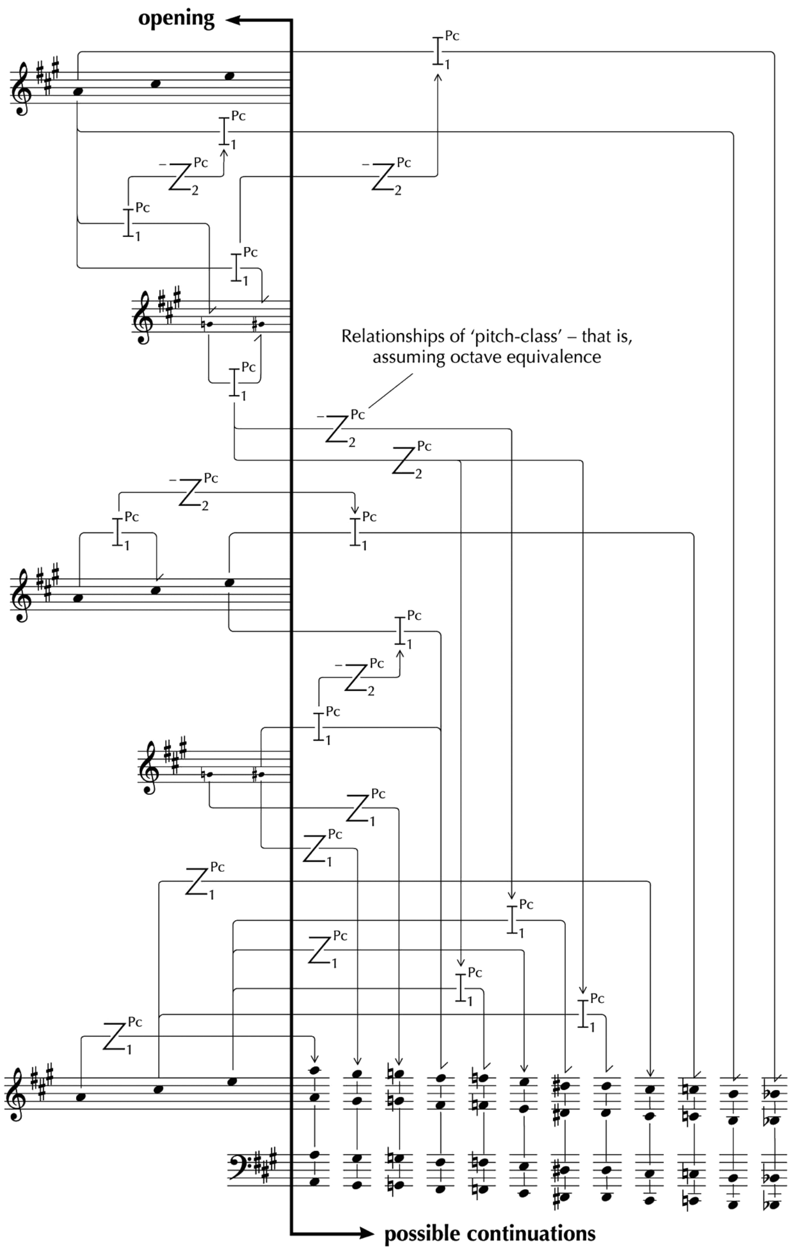


Figure 4. Potential coherent intraopus continuations from the first three notes of the theme and its principal counterpoint in the 3rd movement of Rachmaninov's 2nd Symphony

Now consider **B**, in relation to another listener who is familiar with late-Romantic Western tonality. He will have heard certain interval combinations and pitch transitions many times, and these provide a framework – a schema – that influences his expectation. The schema constrains the implications that result from **A** alone, imbuing the set of expectations that arise intraopusly with greater specificity. At the same time, **A** *grounds* **B** (an abstract set of relationships) at an absolute level of pitch: it offers *local context*. For example, if the listener heard the opening segment (including the accompaniment) as unfolding a tonic minor-seventh chord, suggesting the tonal centre of A major, then the diatonic options for the next melodic note, within the range of an octave, would be as follow (see Figure 5).

The musical validity of this model can be illustrated with the following potential harmonizations; see Figure 6.

But there is more to **B** (schematic expectation) in the domain of pitch than a framework of relative values: qualia are weighted probabilistically according to their frequency of past occurrence (Huron, 2006, pp. 147ff). The more often something has happened in the past, the

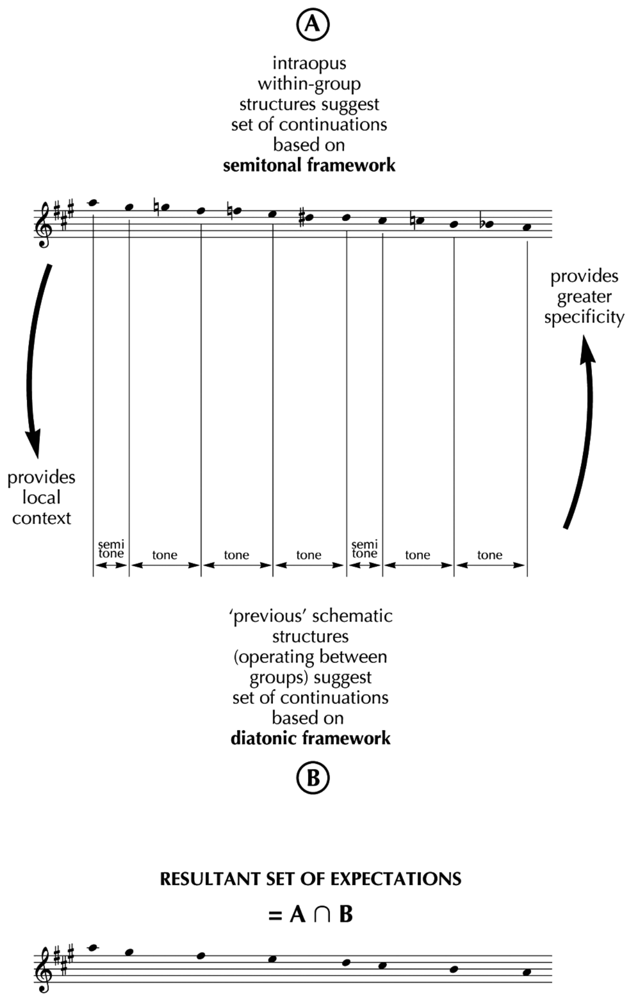


Figure 5. Indicative interaction of 'within group' and 'schematic' structures in expection

Figure 6. Potential coherent continuations following bar 1 of the 3rd Movement of Rachmaninov's 2nd Symphony, based on within-group and schematic structures

stronger, *ceteris paribus*, will be a listener's sense that it will occur again. In the realm of pitch, sub-domains include:

(a) *range*, whereby 'mid-range' pitches are encountered more frequently (and are therefore felt to be more probable) than those at the extremes;
(b) *interval size*, with a tendency for smaller intervals to be used more than larger ones;
(c) *scale degrees*, with context-specific differences in the frequencies of utilization; and
(d) *scale-degree transitions*, which, again, show context-specific patterns of occurrence.

Tonal degree following initial ***i iii v***	*Example from Western classical repertoire*	*Number of instances cited in Barlow & Morgenstern (1948)*	*Relative frequency*
i	Bach: Prelude No. 9 in E Major, BWV 854	50	0.33
ii	Handel: Sonata in D Major, Op. 1, No. 13 for Violin and Continuo; 1st Movement	4	0.03
iii	Schubert: Sonata for Violin and Piano, Op. 137, No. 1; 1st Movement	25	0.17
iv	Mozart: Symphony No. 40 in G Minor, K. 550; Trio	14	0.09
v	Haydn: Symphony No. 104 in D Major; Minuet	31	0.21
vi	Wagner: Parsifal; Overture	18	0.12
vii	Stravinsky: Capriccio for Orchestra (rev. 1949); 3rd Movement	8	0.05
Totals		**150**	**1**

Figure 7. Examples of melodic continuations following the opening i-iii-v in the Western classical tradition

To take a further example: consider the frequencies with which different continuations of Rachmaninov's opening melodic gesture $\hat{1}$–$\hat{2}$–$\hat{3}$ occur within previously established works of the Western classical repertoire, as gauged by Barlow and Morgenstern (1948); see Figure 7. There is no suggestion that the relative frequencies with which this series of scale degrees exist translate directly into probabilities within an expectancy framework, since other factors necessarily play a part, including listeners' degree of familiarity with the works concerned. Nor is it clear how 'scale-degree transition' interacts with the other sub-domains of pitch. However, it is evident that, although **B** accords **A** greater specificity, the effect is still weak, offering listeners a variety of coherent continuations.

The position with regard to **C** – veridical memory traces – is different, however. Such traces offer more or less specific indications of what is to come, depending on how similar the new material is to that heard in the past. **C** adds specificity to the implications deriving from **A** and **B**, which together provide the context in which expectations from **C** are realized.

Consider this assertion in relation to 'current structures'. Take the opening of bar 3 of the Rachmaninov slow movement. In the first violin part, *between-group* projections potentially

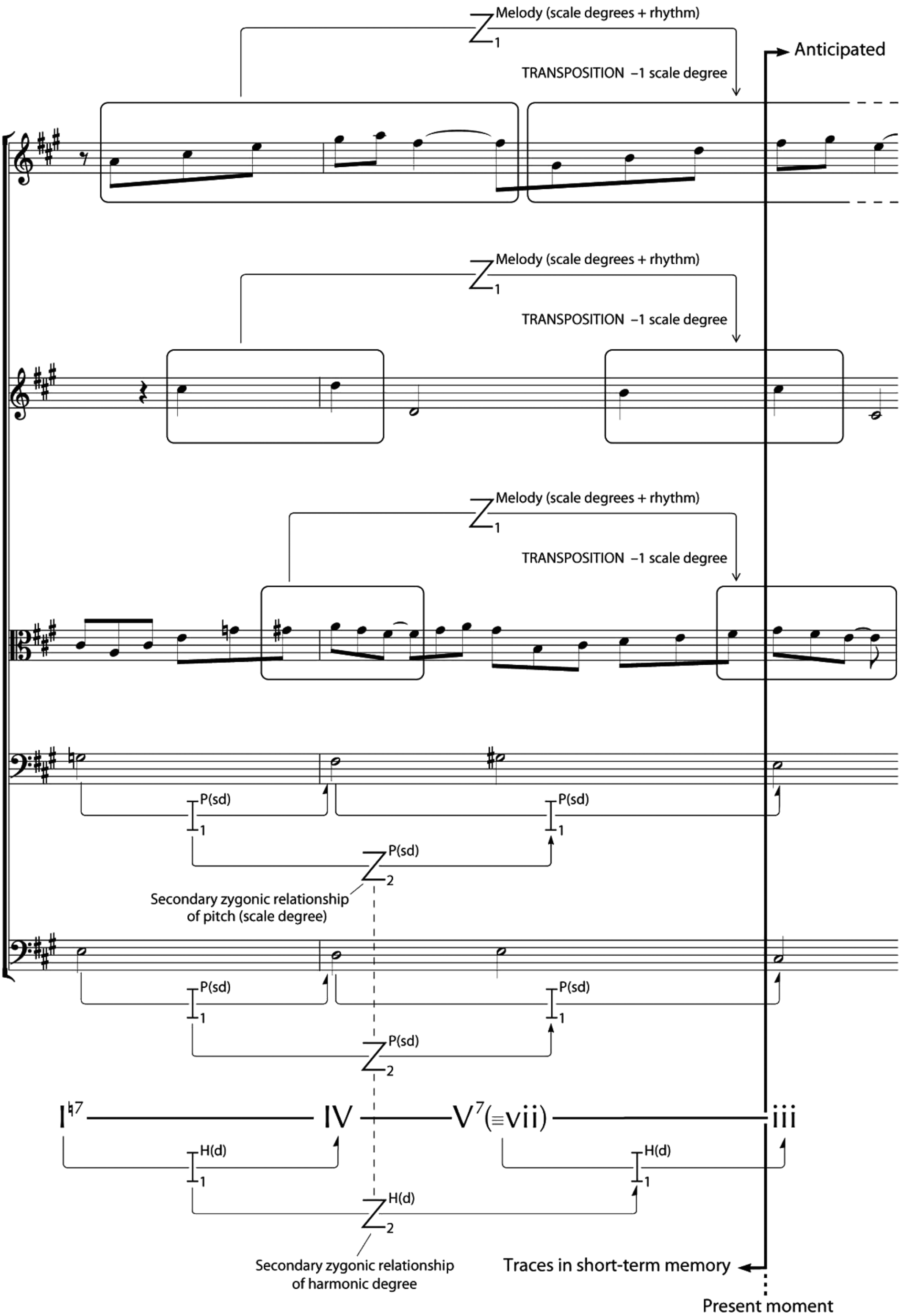


Figure 8. Examples of between-group projections in the 3rd Movement of Rachmaninov's 2nd Symphony

kick in: secure predictions enabled through zygonic invariants (series of relationships operating in parallel). A similar teleological drive may characterize the perception of the inner parts and the bass-line too; see Figure 8.

What is the nature of the interaction between **A**, **B**, and **C** at this point? It is hypothesized that **C** imbues the general tendencies suggested by **A** and **B** with a sense of the specific. At the same time, **A** and **B** provide a context in which the projections arising from **C** are grounded. The nature of this interaction becomes even more telling when expectations arise from veridical memories of previous performances of the piece, which may have involved different tempi, dynamics, and auditory environments. In such circumstances, how veridical is 'veridical'?[2] Here, a version of the 'chameleon effect' (Ockelford, 2005, p. 96) ensures that essence of the musical discourse, as defined by primary relationships of pitch and secondary relationships of onset, is held invariant, while the 'carriers' of the message (absolute pitch, tempo, timbre, and loudness) are modified to ensure congruence with the local musical context (Boulez, 1971, p. 37) (see Figure 9).

The mechanism through which the relative certainties offered by **C** work with the comparative *un*certainties provided by **A** and **B** has been the subject of much music-psychological and aesthetic debate. The zygonic model suggests that the relationship between the two changes according to how familiar listeners are with a piece. At a first hearing, *general* implications (and therefore expectations) stemming from internal patterning – **A** – and stylistic data – **B** – will tend to dominate, with *specific* implications and expectations – **C** – limited to immediate or otherwise memorable repetition or variation of chunks of material, found, for example, in ostinati, sequences, and recapitulations. With subsequent hearings, this balance gradually changes as veridically based expectations come to prevail (see Figure 2 earlier).

Just how the implications inherent in musical structure and the evolving expectations associated with them affect the listening experience is an issue that preoccupied Meyer. His initial proposition was that an affective response would be aroused when an expectation activated by a musical stimulus – a 'tendency to respond' – was inhibited (Meyer, 1956, p. 31). This thesis proved contentious, though. How could one reconcile the uncertainty deemed necessary to stimulate affect with repeated hearings, since people often listen to pieces many times yet continue to enjoy them? Indeed, we typically react most strongly to familiar music (Panskepp, 1995, p. 172). It cannot be the case, though, for a piece one has memorized 'that the ebb and flow of partially fulfilled expectations control one's enjoyment of it: every note is exactly what is expected' (Bever, 1988, p. 166).

Meyer countered arguments like this in various ways (see, for example, Meyer, 1967, pp. 42ff). His final thoughts on the subject (2001) involved the 'willing suspension of disbelief', whereby listeners generate an aesthetic illusion, ignoring their knowledge of a piece and hearing it as if for the first time (2001, p. 352). Some 10 years earlier, Ray Jackendoff had questioned thinking along these lines (1991, pp. 224–228), as it seemed to 'conflate enjoying a piece with not remembering how it goes.' However, Jackendoff proposed 'rescuing' Meyer's expectation theory (p. 228) by suggesting that violations of what is expected may occur on a subconscious level, involving a closed module for music processing. This effectively always hears a piece as if for the first time, thereby ensuring that affect remains intact (cf. Bharucha, 1994, pp. 215–216; Fodor, 1983; Margulis, 2005; Schmuckler, 1989, p. 114). It may be that Meyer's original assertion would be better couched in terms of expectation in music working through the *nonconscious* (rather than the willing) suspension of disbelief. Whatever the precise neurocognitive processes involved, though, the zygonic model of expectation, through its three sources of projection (**A**, **B**, and **C**), suggests that *being able to anticipate what is in stylistic terms the relatively unexpected enables us to relish it all the more* (Ockelford, 2006, p. 134). This accords with Huron's hypothesis (2006) that a key component in the pleasurable experience that music

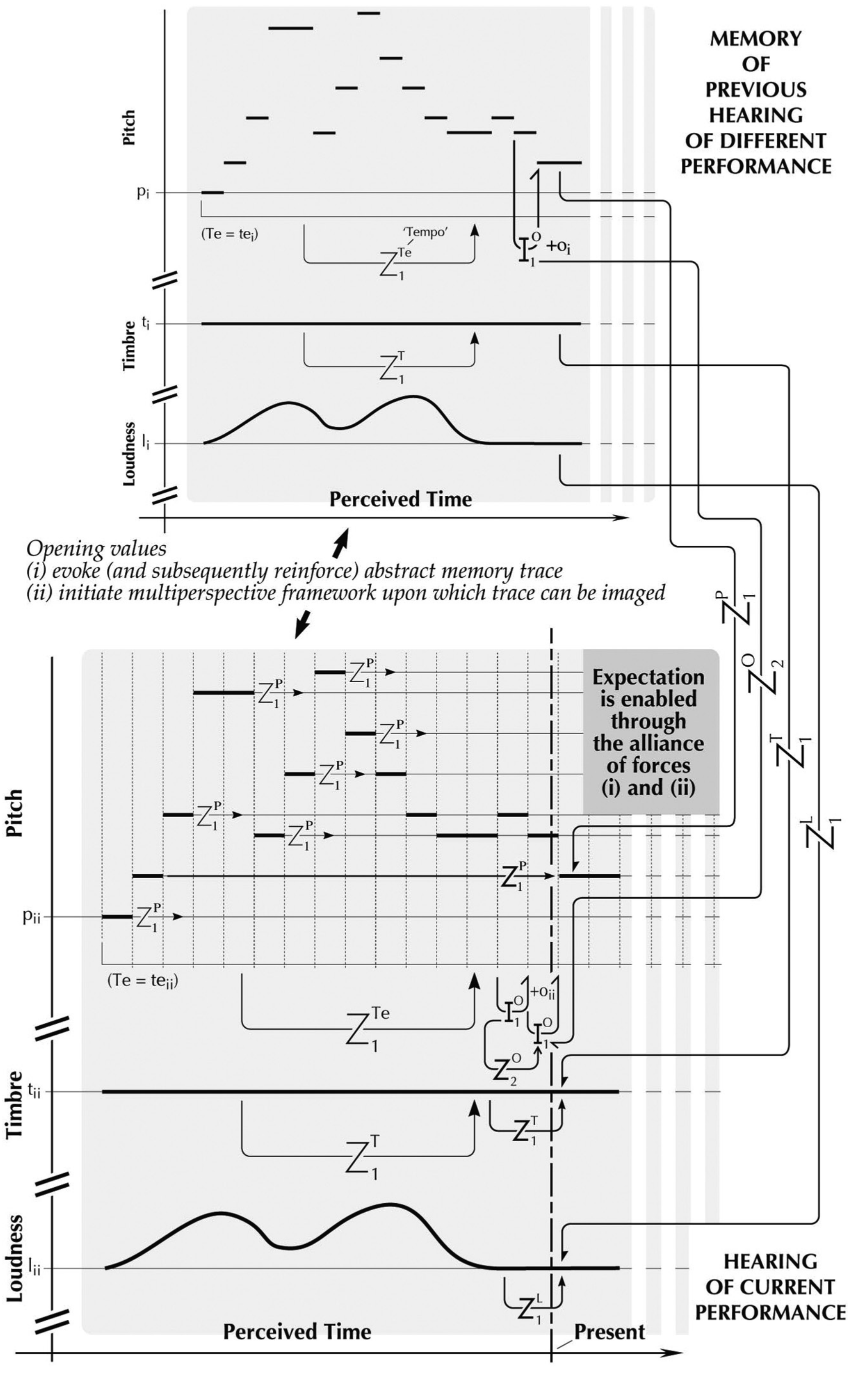


Figure 9. The 'chameleon' effect operating in musical expectation

affords – its 'sweet anticipation' – is the succession of subconscious cognitive rewards that our ability to make correct prognostications offers.

To date, despite its intuitive consonance with the aesthetic experience of listening to music, the zygonic theory of expectation has remained just that – a theory. While it would be difficult to test the whole model empirically at one time, given certain reasonable assumptions, features of it can be used to produce testable predictions. Three of these are reported here.

Designing the empirical work: Rationale, constraints, and assumptions

Our research questions are as follows. With reference to the zygonic model shown in Figure 2:

(a) Is there evidence to support the hypothesis that, *ceteris paribus*, expectations arising from 'current structures, within groups' and 'previous structures, schematically encoded' interact to produce a *general* sense of what may follow?
(b) Is there evidence to support the hypothesis that expectations arising from 'current structures, between groups' produce a *specific* sense of what is to follow?
(c) Is there evidence to support the hypothesis that (b) and (a) interact, whereby (b) adds greater specificity to (a), and (a) grounds (b) in a local context?

These questions engage **A** and **B** and their correlates **1**, **2**, and **3**.

A number of constraints were required to make the empirical work manageable. First, there were restrictions pertaining to the design and utilization of the stimulus.

Constraint 1

The domain in which expectations were predicted and elicited should be *pitch*, since this constitutes the principal structure-bearing dimension of the 'what' in music (working in tandem with the 'when' afforded by patterns of interonset intervals; Boulez, 1971). It further required tessitura, tempo, timbre, and loudness to be as 'neutral' as possible, to avoid potentially confounding effects.

Constraint 2

The major diatonic scale (see Figure 5 earlier) should be used as a framework for the stimulus, and the material should conform to Western tonal 'common practice', with which a broad spectrum of subjects would be familiar. All seven available scale degrees should appear within the span of an octave to facilitate prediction and analysis of the results.

Constraint 3

A single line should be used, to keep the stimulus as simple as possible, avoiding the complexity of expectations in one part potentially influencing those in another.

Constraint 4

New material should be created for the stimulus, to avoid the danger that **C** (previous structures encoded veridically) would figure in subjects' responses.

Constraint 5

For similar reasons, subjects should be limited to one hearing.

Constraint 6

To gauge the impact of between-group structures (an element of **A**), the melody should contain clusters of notes that are repeated or transposed, as well as having episodes in which no such connections exist.

Constraint 7

Given the method of data collection that was adopted (where participants sing the note they expect to come next – see below), and given that both adult males and females were involved, the melody should be positioned within two alternative pitch spans that were mid-range for 'typical' men's and women's voices.

Working within these constraints, the following stimulus was created (Figure 10).

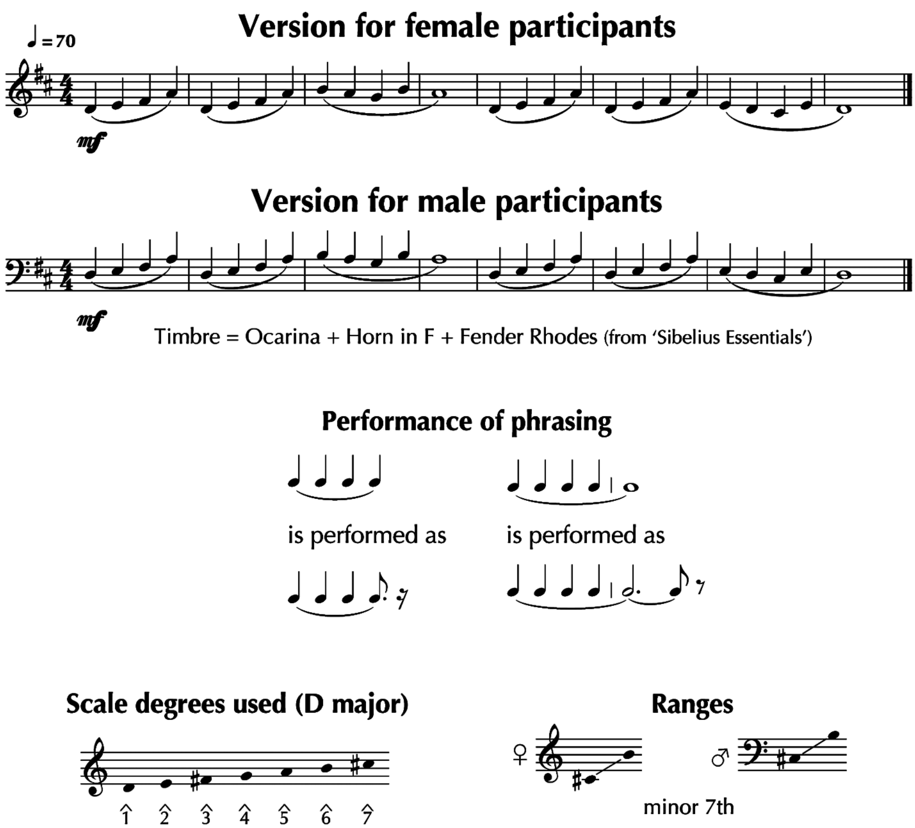


Figure 10. Stimuli used, and their characteristics

In accordance with Constraint 1, the pace is moderate, with little rhythmic variety (the two significantly longer notes serving to underscore the ends of phrases); timbre is unvarying and rich (though intended to be stylistically 'non-specific'); and most musical information is conveyed in the domain of pitch. Each scale-degree occurs at least once, in the key of D major, over the range of a minor 7th (Constraint 2), situated in the 3rd octave for men and the 4th octave for women (Constraint 7). By the fifth note, the unaccompanied melody (Constraint 3) does not conform to any well-known tunes in the Western classical or popular repertories (Constraint 4).

The melody is in the form a_1 a_1 b_1 a_1 a_1 b_2 (Constraint 6), which can be summarized as A_1 A_2 (see Figure 11). As we are seeking to determine how the perception of between-group

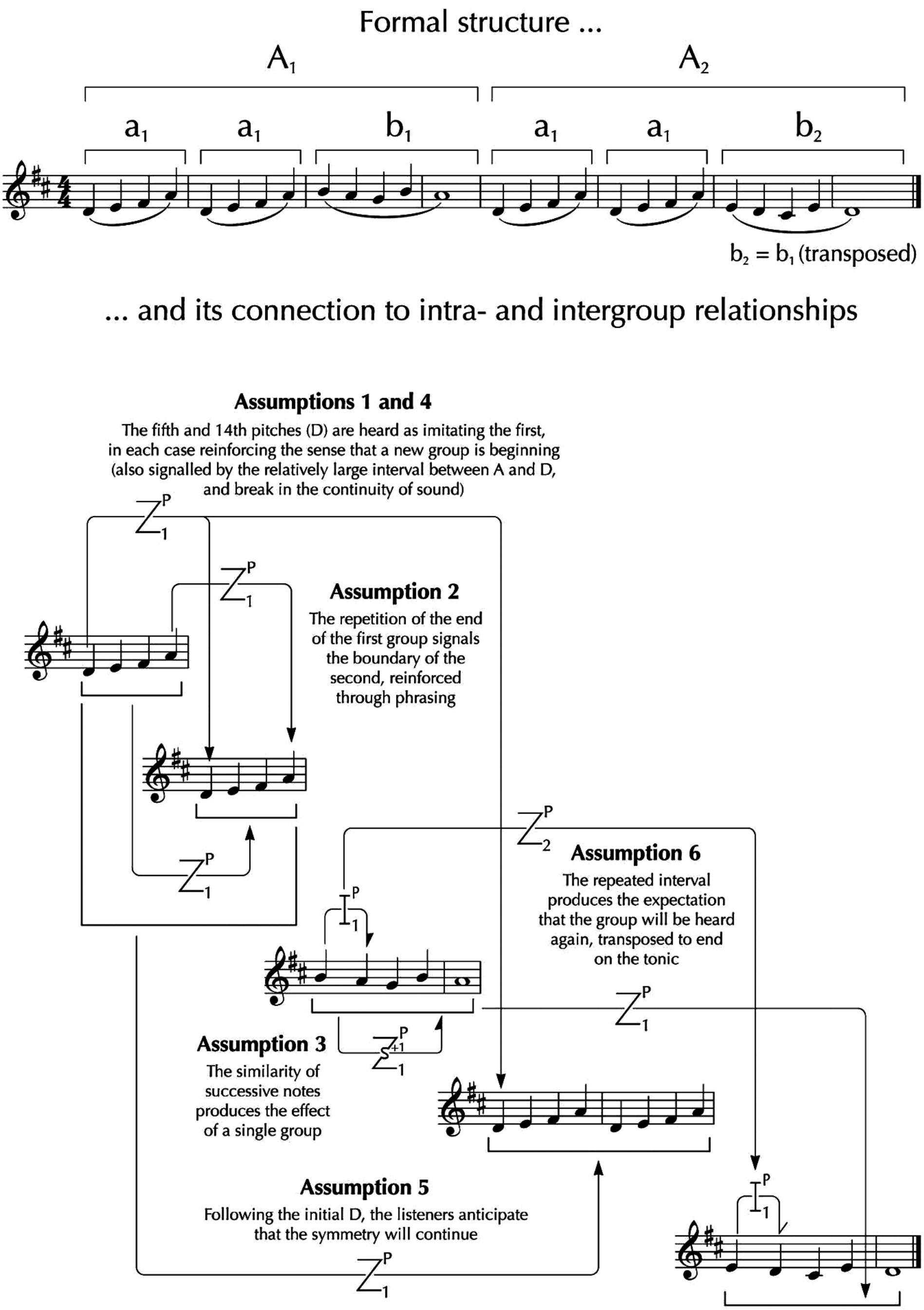


Figure 11. Stimulus melody and musicological and psychomusicological analyses

relationships impacts on expectation, it is crucial to gauge the detectability of the structure, albeit nonconsciously, in first-time listening. To this end, a psychomusicological analysis will be undertaken, using zygonic theory. This draws on the intuitions of the analyst – here, the same as the composer – and is therefore susceptible to problems of subjectivity. However, for the empirical work to get off the ground, a predictive model is essential, making certain assumptions inevitable. These could subsequently be modified, if necessary, in the light of the results obtained from a range of listeners, hearing the stimulus with no structural preconceptions. The assumptions are listed below and illustrated in Figure 11.

Assumption 1

The boundary between the first group a_1 and its repetition will be detected as the second D is heard, on account of all or any of three signals: the shortening of the preceding A, which leaves a discernible gap in the continuity of sound (see Lerdahl & Jackendoff's grouping preference rules, 1983, pp. 43ff); the interval of the descending 5th between the A and D, which is comparatively large, providing a relative melodic discontinuity (Bregman, 1990, pp. 461ff); and the fact the D is a repetition of the opening pitch – potentially an indication that the first motif (or a variant of it) is about to restart (see Lerdahl & Jackendoff's notion of 'parallelism', 1983, p. 51). That is, it is hypothesized that schemata pertaining to the way in which formal structures typically unfold (an aspect of **B**) may play a part in determining group boundaries.

Assumption 2

The second A (at the end of bar 2) will be heard as concluding the second group a1, due to the equivalent note in the first appearance of a_1 fulfilling that function (Lerdahl & Jackendoff, 1983). It is assumed that structural cognition will be reinforced on hearing the B that opens bar 3, through the short break in sound that occurs before it.

Assumption 3

b_1 (bars 3 and 4) will be heard as single group, due to the adjacency of successive pitches, the continuity of sound, and the longer duration of the final note which, it is expected, will strongly signal a phrase boundary, subsequently reinforced upon hearing the brief gap in sound before the next note, D.[3]

Assumption 4

This (third) D will indicate a return to the opening segment A_1 (or a variant of it), due to the symmetry that is implied (through listeners' assumed experience of archetypal formal structures) by a repeat of the opening note after the end of the first phrase. It is anticipated that the perception of metre will be established by this point (bolstered by the recognition of previous grouping structures, with their strong metrical correspondence), and that listeners will suppose that the initial 4/4 will continue, running in parallel with any motivic repetition and transformation, thereby reinforcing expectation in the domain of pitch (cf. Temperley, 1995, p. 141; Ockelford, 2009, p. 75).

Assumption 5

Listeners will anticipate that the symmetry continues, with $a_1 + a_1$ followed by b_1 or a variant of it (which, in simple tonal melodies, would typically resolve the dominant that was heard at the halfway point – A – onto the tonic, D).

Assumption 6

The E at the beginning of bar 7 (after the second reprise of a_1) will register with listeners as an indication that something different is about to happen. However, it is not until the onset of the following D (which, like the opening of bar 3, frames a descending major 2nd) that listeners are expected to anticipate a transposed version of b_1, concluding on the tonic.

Further constraints and assumptions, arising partly from the design of the melody, impinge upon the way in which the patterns of expectation that evolve as the stimulus is heard are modelled according to zygonic theory: limitations and suppositions that will enable quanta to be assigned to the anticipated relative strengths of expectation involved.

Constraint 8

In relation to **A1** ('current structures, within groups', see Figure 2 earlier), since little is known about how expectations arising from pitches or intervals interact (from primary and secondary zygonic relationships respectively), it was decided, for this initial investigation, to model only the impact of perfect and imperfect *primary* zygons that may conceivably operate from a given note (see Figure 12).

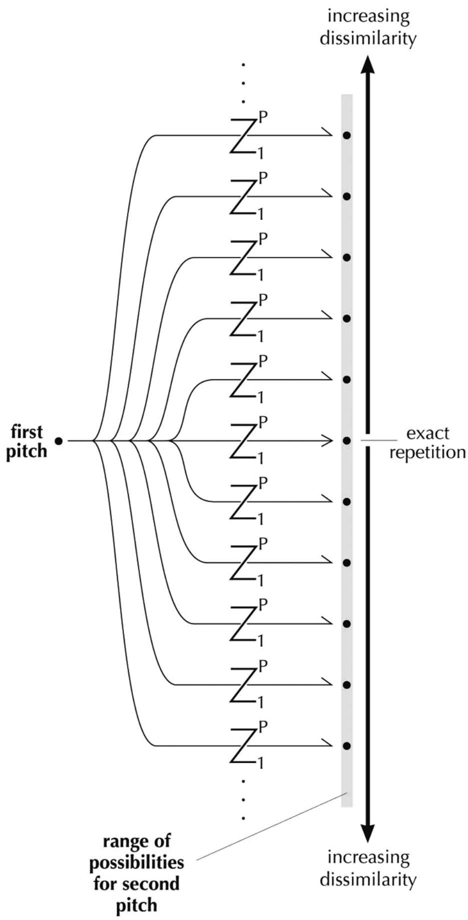


Figure 12. Hypothetical range of potential zygonic relationships from a given pitch

Assumption 7

Such relationships will be subject to an *adjacency* effect, since zygonic theory suggests that, *ceteris paribus*, the strength of expectation will be inversely proportional to the dissimilarity of a predicted note to the stimulus, although the precise nature of the relationship is not determined (cf. Ortmann, 1926, p. 30). Turning to the empirical data pertaining to the melodic intervals occurring between successive notes that are available from Ortmann (1926) and Huron (2001, p. 25), the statistical picture is actually more complex than this: exact repetition tends to be used less frequently than intervals of a second (the desire for similarity apparently outweighing the wish for duplication), and the interval of an octave arises more often than sixths and sevenths (arguably due to the influence of the harmonic series; see Ortmann, 1926, p. 31). Combining the findings of the two studies yields intervallic data relating to over 340 pieces from 10 cultures, which can be expressed in terms of differences between scale steps as follows (see Figure 13).

These data will be used to form what Huron would term a 'serviceable heuristic', to quantify the expectations pertaining to **B2** ('previous structures, between groups, encoded schematically'). This heuristic is limited, since it reflects intervals from a relatively small number of pieces and extends only an octave in each direction from the target pitch. Moreover, the data do not take into account context-specific features, such as the scale-steps involved, the position of a given interval in a particular melody, or its direction (ascending or descending). However, in terms of building a general model, we will assume that these surface details can reasonably be set aside for the time being (and subsequently revisited in the light of empirical findings): the important issue is to identify trends with which the data accord and which are likely to be intuitively apprehensible.

Given that the overall probability must total 1, and on the working assumption that the heuristic is bidirectional and operates symmetrically about the pitch that is presented, the model predicts that, following the first note of the stimulus melody, pitches will be anticipated with the following probabilities (see Figure 14).

Interval	Ortmann/Huron relative frequency
0 scale-steps (unison)	0.24
±1 scale-step (2nd)	0.42
±2 scale-steps (3rd)	0.17
±3 scale-steps (4th)	0.08
±4 scale-steps (5th)	0.03
±5 scale-steps (6th)	0.02
±6 scale-steps (7th)	0.01
±7 scale-steps (octave)	0.03
Total	**1.00**

Figure 13. Intervallic probabilities, after Ortmann (1926) and Huron (2001)

Current pitch	Next pitch	Interval	Predicted probability
	D5	octave up	0.02
	C♯5	7th up	0.01
	B4	6thup	0.01
	A4	5th up	0.02
	G4	4th up	0.05
	F♯4	3rd up	0.10
	E4	2nd up	0.24
D4	D4	unison	0.14
	C♯4	2nd down	0.24
	B3	3rd down	0.10
	A3	4th down	0.05
	G3	5th down	0.02
	F♯3	6th down	0.01
	E3	7th down	0.01
	D3	octave down	0.02
		Total	**1**

Figure 14. Predicted probabilities of expectation using the adjacency model following the first note of the melody

Assumption 8

The significance of adjacency in prediction may extend to stimuli other than the one heard most recently (Ockelford, 2006, p. 108), and a *recency* effect is postulated, whereby the closer the stimulus to the point at which expectation occurs, the greater its impact on anticipating what will happen next. This is represented schematically in Figure 15.

To enable this model to function predictively – to indicate which pitches are likely to be expected, and with what probabilities – it is necessary to quantify two factors: the *number* of notes (how far back in the sequence to extend) and the *nature* of the relationship between events in terms of their relative impact (for example, a more contemporaneous note could be deemed to exert twice the effect of the one preceding, and so on).

Two criteria were used in the process of quantification: the psychological constraints of working memory, and the mathematical consideration that some of the proportions involved in predictions from just one note are already very small, meaning that significant reduction would render them trivial. Given these two restrictions, a model was devised that took into account a maximum of *four* events, with a *linear* decrement of impact. The latter is calculated such that the predicted probability of an *n*th pitch (p_n) (occurring after n–1 events) is given by the equation:

$$P(p_n) = ({}^{n-1}/_{\Sigma(n-1)})\bullet P(p_{n-1}) + ({}^{n-2}/_{\Sigma(n-1)})\bullet P(p_{n-2}) + \ldots ({}^{1}/_{\Sigma(n-1)})\bullet P(p_1)$$

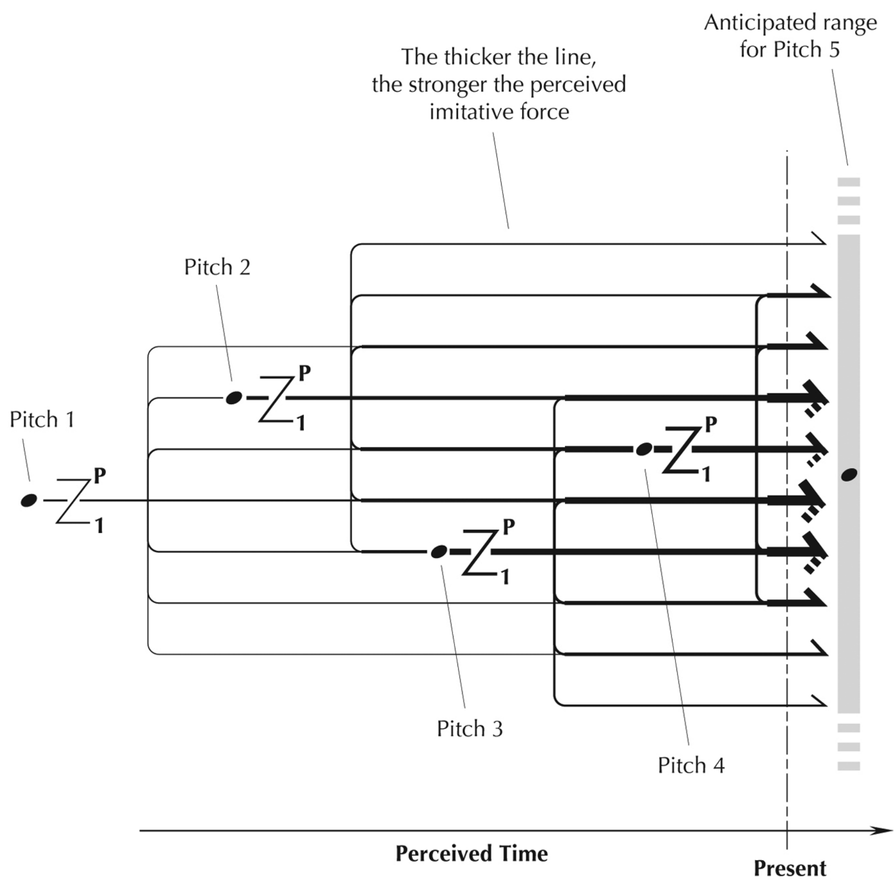


Figure 15. Schematic representation of zygonic adjacency + recency model of expectation in music

To illustrate this principle in action, here are the predicted probabilities of notes 3, 4, and 5 from the stimulus melody (Figure 16).[4]

Assumption 9

In relation to **A3** of the zygonic model ('current structures, between-groups'), groups will be recognized from the first note onwards in the case of exact repetition (see Assumptions 1 and 4, and Figure 11 above) and from the first interval (that is, the second note) in the case of transposition.

Assumption 10

The combined 'adjacency/recency' effect (**A1** + **B2**) will interact with the 'between-groups' effect (**A3**), whereby the strength of expectation generated by a group of repeated or transposed notes will increase rapidly as the sequence of pitches and intervals is heard again. That is, as listeners become more certain that what they are hearing is familiar, it is assumed they will make increasingly specific predictions as to what is likely to occur next. No empirical data are available to quantify this conjecture, but it is postulated that the impact ratio between the two factors – (**A1** + **B2**) : **A3** – will change

Following the first two notes

1st pitch	Probabilities of next pitch based on adjacency only	Probabilities based on adjacency + recency	2nd pitch	Probabilities of next pitch based on adjacency only	Probabilities based on adjacency + recency	3rd pitch	Probabilities based on pitches 1 & 2
			E5	0.02	**0.01**	E5	**0.01**
D5	0.02	**0.01**	D5	0.01	**0.00**	D5	**0.01**
C 5	0.01	**0.00**	C 5	0.01	**0.01**	C 5	**0.01**
B4	0.01	**0.00**	B4	0.02	**0.01**	B4	**0.02**
A4	0.02	**0.01**	A4	0.05	**0.03**	A4	**0.04**
G4	0.05	**0.02**	G4	0.10	**0.06**	G4	**0.08**
F 4	0.10	**0.03**	F 4	0.24	**0.16**	F 4	**0.19**
E4	0.24	**0.08**	E4	0.14	**0.09**	E4	**0.17**
D4	0.14	**0.05**	D4	0.24	**0.16**	D4	**0.20**
C 4	0.24	**0.08**	C 4	0.10	**0.06**	C 4	**0.14**
B3	0.10	**0.03**	B3	0.05	**0.03**	B3	**0.06**
A3	0.05	**0.02**	A3	0.02	**0.01**	A3	**0.03**
G3	0.02	**0.01**	G3	0.01	**0.01**	G3	**0.01**
F 3	0.01	**0.00**	F 3	0.01	**0.00**	F 3	**0.01**
E3	0.01	**0.00**	E3	0.02	**0.01**	E3	**0.01**
D3	0.02	**0.01**				D3	**0.01**
		x0.33			**x0.67**		**1.00**

Following the first three notes

1st pitch	Probabilities of next pitch based on adjacency only	Probabilities based on adjacency + recency	2nd pitch	Probabilities of next pitch based on adjacency only	Probabilities based on adjacency + recency	3rd pitch	Probabilities of next pitch based on adjacency only	Probabilities based on adjacency + recency	4th pitch	Probabilities based on pitches 1–3
						F 5	0.02	**0.01**	F 5	**0.01**
			E5	0.02	**0.01**	E5	0.01	**0.00**	E5	**0.01**
D5	0.02	**0.00**	D5	0.01	**0.00**	D5	0.01	**0.01**	D5	**0.01**
C 5	0.01	**0.00**	C 5	0.01	**0.00**	C 5	0.02	**0.01**	C 5	**0.01**
B4	0.01	**0.00**	B4	0.02	**0.01**	B4	0.05	**0.02**	B4	**0.03**
A4	0.02	**0.00**	A4	0.05	**0.02**	A4	0.10	**0.05**	A4	**0.07**
G4	0.05	**0.01**	G4	0.10	**0.03**	G4	0.24	**0.12**	G4	**0.16**
F 4	0.10	**0.02**	F 4	0.24	**0.08**	F 4	0.14	**0.07**	F 4	**0.16**
E4	0.24	**0.04**	E4	0.14	**0.05**	E4	0.24	**0.12**	E4	**0.20**
D4	0.14	**0.02**	D4	0.24	**0.08**	D4	0.10	**0.05**	D4	**0.15**
C 4	0.24	**0.04**	C 4	0.10	**0.03**	C 4	0.05	**0.02**	C 4	**0.10**
B3	0.10	**0.02**	B3	0.05	**0.02**	B3	0.02	**0.01**	B3	**0.04**
A3	0.05	**0.01**	A3	0.02	**0.01**	A3	0.01	**0.01**	A3	**0.02**
G3	0.02	**0.00**	G3	0.01	**0.00**	G3	0.01	**0.00**	G3	**0.01**
F 3	0.01	**0.00**	F 3	0.01	**0.00**	F 3	0.02	**0.01**	F 3	**0.01**
E3	0.01	**0.00**	E3	0.02	**0.01**				E3	**0.01**
D3	0.02	**0.00**							D3	**0.00**
		x0.17			**x0.33**			**x0.50**		**1.00**

Following the first four notes

1st pitch	Probabilities of next pitch based on adjacency only	Probabilities based on adjacency + recency	2nd pitch	Probabilities of next pitch based on adjacency only	Probabilities based on adjacency + recency	3rd pitch	Probabilities of next pitch based on adjacency only	Probabilities based on adjacency + recency	4th pitch	Probabilities of next pitch based on adjacency only	Probabilities based on adjacency + recency	5th pitch	Probabilities based on pitches 1–4
									A5	0.02	**0.01**	A5	**0.01**
									G5	0.01	**0.00**	G5	**0.00**
						F 5	0.02	**0.01**	F 5	0.01	**0.00**	F 5	**0.01**
			E5	0.02	**0.00**	E5	0.01	**0.00**	E5	0.02	**0.01**	E5	**0.01**
D5	0.02	**0.00**	D5	0.01	**0.00**	D5	0.01	**0.00**	D5	0.05	**0.02**	D5	**0.02**
C 5	0.01	**0.00**	C 5	0.01	**0.00**	C 5	0.02	**0.01**	C 5	0.10	**0.04**	C 5	**0.05**
B4	0.01	**0.00**	B4	0.02	**0.00**	B4	0.05	**0.01**	B4	0.24	**0.10**	B4	**0.11**
A4	0.02	**0.00**	A4	0.05	**0.01**	A4	0.10	**0.03**	A4	0.14	**0.05**	A4	**0.09**
G4	0.05	**0.00**	G4	0.10	**0.02**	G4	0.24	**0.07**	G4	0.24	**0.10**	G4	**0.19**
F 4	0.10	**0.01**	F 4	0.24	**0.05**	F 4	0.14	**0.04**	F 4	0.10	**0.04**	F 4	**0.14**
E4	0.24	**0.02**	E4	0.14	**0.03**	E4	0.24	**0.07**	E4	0.05	**0.02**	E4	**0.14**
D4	0.14	**0.01**	D4	0.24	**0.05**	D4	0.10	**0.03**	D4	0.02	**0.01**	D4	**0.10**
C 4	0.24	**0.02**	C 4	0.10	**0.02**	C 4	0.05	**0.01**	C 4	0.01	**0.00**	C 4	**0.06**
B3	0.10	**0.01**	B3	0.05	**0.01**	B3	0.02	**0.01**	B3	0.01	**0.00**	B3	**0.03**
A3	0.05	**0.00**	A3	0.02	**0.00**	A3	0.01	**0.00**	A3	0.02	**0.01**	A3	**0.02**
G3	0.02	**0.00**	G3	0.01	**0.00**	G3	0.01	**0.00**				G3	**0.01**
F 3	0.01	**0.00**	F 3	0.01	**0.00**	F 3	0.02	**0.01**				F 3	**0.01**
E3	0.01	**0.00**	E3	0.02	**0.00**							E3	**0.00**
D3	0.02	**0.00**										D3	**0.00**
		x0.10			**x0.20**			**x0.30**			**x0.40**		**1.00**

Figure 16. Predicted probabilities of expectation using the adjacency + recency model following the first 2, 3, and 4 notes of the melody

If group is a repetition	If group is a transposition	'Within-group' influence	'Between-group' influence
After note 1	After note 2	0.5	0.5
After note 2	After note 3	0.25	0.75
After note 3	After note 4	0.125	0.875
After note 4	After note 5	0.0625	0.9375
After note 5	After note 6	0.0313	0.9688
After note 6	After note 7	0.0156	0.9844

Figure 17. Predicted ratios between 'within-group' and 'between-group' influence

exponentially as one proceeds further through the group that is repeated or transposed, such that, after the *n*th pitch, the 'between-group' and 'adjacency/recency' influences will be deemed to be $1 - {}^{1}/_{2^n}$ and ${}^{1}/_{2^n}$ respectively in the case of repetition, a ratio of $(2^n - 1) : 1$, and, in the case of transposition, $1 - {}^{1}/_{2^{(n-1)}}$ and ${}^{1}/_{2^{(n-1)}}$, a ratio of $(2^{n-1} - 1) : 1$. Figure 17 shows the ratios in relation to notes 1–6 of a group.

Figure 18 illustrates these equations operating in relation to notes 5, 6, 7, and 8 of the stimulus.

Although the principles driving the model are straightforward, its quantification results in a plethora of data, whose manipulation is time-consuming and arguably of little perceptual consequence at the periphery of the action (where values are very small). For example, after only four notes of a repeated group, the relative impact of adjacency/recency is negligible at the extremes of pitch, although the model still sets out the calculations. A certain degree of over-specificity is probably inevitable in the early stages of developing any such protocol that seeks to emulate complex human behaviour, when key parameters are still being determined. However, it is hoped that future empirical work may suggest refinements and ways in which the model may be simplified without losing its flexibility or anticipated predictive power.

Hence we have three models of expectation rooted in zygonic theory that will enable us to tackle the research questions: the one-factor 'adjacency' model, '**Z1**'; the two-factor 'adjacency/recency' model, '**Z2**'; and the three-factor 'adjacency/recency/between-groups' model, '**Z3**'. The next issue is a methodological one: how to elicit and record responses with which the predictions can be compared. There are a number of ways of testing musical expectation (Huron, 2006, pp. 42–52). Three could potentially produce data suitable for comparison with the zygonic models' outputs: the 'progressive probe-tone method' (Krumhansl, 1990, pp. 214ff); the 'betting paradigm' (Huron, 2006, pp. 47–9); and the 'production method' – either through singing (Carlsen, Divenyi, & Taylor, 1970) or playing (Povel, 1996; Schmuckler, 1989). The first two approaches produce a *range* of responses for every listener in relation to each note of the stimulus: probabilities are generated for a set of pitches that indicate each one's perceived suitability as a potential successor to those that are presented. This has the advantage that each subject produces a full set of data that can be compared with the multiple, probabilistic continuations projected by the zygonic models. However, this benefit is outweighed by problems associated with the progressive probe-tone method and the betting paradigm, which are very time-consuming to administer and, in the case of the former, restrict subjects to a limited number of continuations to which to react, and in the latter, constrain the choice of subjects to those with formal music education.

Following the first five notes

2nd pitch	Probabilities based on adjacency + recency	3rd pitch	Probabilities based on adjacency + recency	4th pitch	Probabilities based on adjacency + recency	5th pitch	Probabilities based on adjacency + recency	6th pitch	Probabilities based on pitches 2–5	x0.5 = 'within-group' factor	'Between-group' factor	**Predicted probability of expectation**
				A5	0.01			A5	0.01	**0.00**		**0.00**
				G5	0.00			G5	0.00	**0.00**		**0.00**
		F 5	0.00	F 5	0.00			F 5	0.01	**0.00**		**0.00**
E5	0.00	E5	0.00	E5	0.01			E5	0.01	**0.00**		**0.00**
D5	0.00	D5	0.00	D5	0.01	D5	0.01	D5	0.02	**0.01**		**0.01**
C 5	0.00	C 5	0.00	C 5	0.03	C 5	0.00	C 5	0.04	**0.02**		**0.02**
B4	0.00	B4	0.01	B4	0.07	B4	0.00	B4	0.09	**0.04**		**0.04**
A4	0.00	A4	0.02	A4	0.04	A4	0.01	A4	0.07	**0.04**		**0.04**
G4	0.01	G4	0.05	G4	0.07	G4	0.02	G4	0.15	**0.07**		**0.07**
F 4	0.02	F 4	0.03	F 4	0.03	F 4	0.04	F 4	0.12	**0.06**		**0.06**
E4	0.01	E4	0.05	E4	0.01	E4	0.10	E4	0.17	**0.09**	0.50	**0.59**
D4	0.02	D4	0.02	D4	0.01	D4	0.05	D4	0.10	**0.05**		**0.05**
C 4	0.01	C 4	0.01	C 4	0.00	C 4	0.10	C 4	0.12	**0.06**		**0.06**
B3	0.00	B3	0.00	B3	0.00	B3	0.04	B3	0.05	**0.02**		**0.02**
A3	0.00	A3	0.00	A3	0.01	A3	0.02	A3	0.03	**0.01**		**0.01**
G3	0.00	G3	0.00			G3	0.01	G3	0.01	**0.00**		**0.00**
F 3	0.00	F 3	0.00			F 3	0.00	F 3	0.01	**0.00**		**0.00**
E3	0.00					E3	0.00	E3	0.00	**0.00**		**0.00**
						D3	0.01	D3	0.01	**0.00**		**0.00**
	0.10		0.20		0.30		0.40		1.00	**0.50**		**1.00**

Following the first six notes

7th pitch	Probabilities based on pitches 3–6	x0.25 = 'within-group' factor	'Between-group' factor	Predicted probability of expectation
A5	0.00	**0.00**		**0.00**
G5	0.00	**0.00**		**0.00**
F 5	0.00	**0.00**		**0.00**
E5	0.01	**0.00**		**0.00**
D5	0.02	**0.00**		**0.00**
C 5	0.03	**0.01**		**0.01**
B4	0.06	**0.02**		**0.02**
A4	0.06	**0.02**		**0.02**
G4	0.12	**0.03**		**0.03**
F 4	0.16	**0.04**	0.75	**0.79**
E4	0.16	**0.04**		**0.04**
D4	0.15	**0.04**		**0.04**
C 4	0.12	**0.03**		**0.03**
B3	0.05	**0.01**		**0.01**
A3	0.03	**0.01**		**0.01**
G3	0.01	**0.00**		**0.00**
F 3	0.01	**0.00**		**0.00**
E3	0.01	**0.00**		**0.00**
D3	0.01	**0.00**		**0.00**
	1.00	**0.25**		**1.00**

Following the first seven notes

8th pitch	Probabilities based on pitches 4–7	x0.13 = 'within-group' factor	'Between-group' factor	Predicted probability of expectation
A5	0.00	**0.00**		**0.00**
G5	0.00	**0.00**		**0.00**
F 5	0.01	**0.00**		**0.00**
E5	0.01	**0.00**		**0.00**
D5	0.01	**0.00**		**0.00**
C 5	0.02	**0.00**		**0.00**
B4	0.05	**0.01**		**0.01**
A4	0.07	**0.01**	0.88	**0.88**
G4	0.16	**0.02**		**0.02**
F 4	0.16	**0.02**		**0.02**
E4	0.19	**0.02**		**0.02**
D4	0.14	**0.02**		**0.02**
C 4	0.10	**0.01**		**0.01**
B3	0.04	**0.01**		**0.01**
A3	0.02	**0.00**		**0.00**
G3	0.01	**0.00**		**0.00**
F 3	0.01	**0.00**		**0.00**
E3	0.01	**0.00**		**0.00**
D3	0.00	**0.00**		**0.00**
	1.00	**0.13**		**1.00**

Following the first eight notes

9th pitch	Probabilities based on pitches 5–8	x0.06 = 'within-group' factor	'Between-group' factor	Predicted probability of expectation
A5	0.01	**0.00**		**0.00**
G5	0.00	**0.00**		**0.00**
F 5	0.01	**0.00**		**0.00**
E5	0.01	**0.00**		**0.00**
D5	0.02	**0.00**		**0.00**
C 5	0.05	**0.00**		**0.00**
B4	0.11	**0.01**		**0.01**
A4	0.09	**0.01**		**0.01**
G4	0.19	**0.01**		**0.01**
F 4	0.14	**0.01**		**0.01**
E4	0.14	**0.01**	0.94	**0.94**
D4	0.10	**0.00**		**0.00**
C 4	0.06	**0.00**		**0.00**
B3	0.03	**0.00**		**0.00**
A3	0.02	**0.00**		**0.00**
G3	0.01	**0.00**		**0.00**
F 3	0.01	**0.00**		**0.00**
E3	0.00	**0.00**		**0.00**
D3	0.00	**0.00**		**0.00**
	1.00	**0.06**		**1.00**

Figure 18. Predicted probabilities of expectation using the adjacency + recency + intergroup model following notes 5, 6, 7, and 8

Both the production methods – singing and playing – mean that subjects can produce only one response per note. Hence, in order to obtain a range of perceived probabilities pertaining to a number of potential future pitches, the responses of many subjects have to be amalgamated. Therefore, if this method is to be used, a further assumption is required.

Assumption 11

A valid impression of human musical expectation in relation to the differing probabilities that potentially pertain to different melodic continuations can be obtained by combining a number of listeners' responses.

A problem with the 'playing' production method is that subjects need to be able to play by ear, restricting the pool of potential subjects, and conceivably resulting in an atypical sample. The potential challenge of having subjects sing responses is that these may be constrained by the tessitura of the voice, and, in the absence of vocal training, that pitches may be produced inaccurately. The first difficulty can be obviated by avoiding registral extremes (Constraint 7), and the second can be countered by using frequency detection software and making reasonable assumptions as to what was intended. However, a number of issues remain: the task is artificial, since listeners do not usually listen to music in incrementally increasing chunks; they have to sing notes, consciously reflecting on and extrapolating from the listening experience in a way that is unnatural; and the pitches they produce – particularly if they prove to be incorrect – may be distracting. Hence, a final assumption is necessary.

Assumption 12

A singing protocol of the type described will not interfere with subjects' listening experiences to such an extent that the responses they offer fail to present a reasonable picture of the expectations that would otherwise occur.

Method

Research participants

Forty subjects, 24 female and 16 male, aged between 21 and 76, mean 34 years, were recruited through direct contact and posters at Roehampton University, London and other community sites in the area. One subject reported minor hearing loss, but this did not appear to interfere with his ability to take part in the experiment and his contribution was included. Subjects were recruited without regard to their musical or, specifically, vocal training. However, only 11 (28%) reported having had no previous formal music education. Of the remaining 29, five (17%) had had less than 2 years specialist input, six (21%) had had 2 to 4 years, nine (31%) had had 4 to 8 years, and nine (31%) had had more than 8 years. Seventeen (59%) also reported having had voice or singing lessons. All subjects reported listening to music every day (15% less than 30 minutes, 45% between 30 and 60 minutes, and 40% more than 60 minutes) and had had significant exposure to Western mainstream pieces. Three also reported listening regularly to other styles, including Indian ragas and Slovakian folk music.

Materials

The stimulus used was the melody shown in Figure 10, in different ranges for males and females. The timbre was instrumentally and stylistically non-specific (to avoid experience of specific instruments and composers' use of them affecting subjects' expectations) yet rich in harmonics and musically 'realistic' (to make the task as ecologically valid as possible), and, to this end, three sounds were blended from the Sibelius 5 software: the 'ocarina', 'horn in F', and 'flute'.

Since listeners were asked to do something that was outside their experience, another – very short – practice melody was created. This was similar to the main stimulus, yet differed from it, so that when the principal melody was subsequently heard, it would be regarded as a distinct musical entity, and 'previous structures', 'between-groups', 'veridical' memories would not be invoked (**C3**). Hence a musical fragment was constructed in a different key (C major) with a contrasting melodic contour. The version for women is shown in Figure 19 (for men this was an octave lower).

To set the auditory scene in each case, two introductions were produced, using the relevant diatonic major scale ascending and descending. This bi-directionality was again intended to counter any 'previous structures', 'between-groups' effect – reinforced by using a different timbre (the piano); see Figure 20.

Environment and apparatus

Data were collected in a soundproofed room, with only the first author and the subject present, to minimize any discomfort that may have been felt in having to sing in front of someone else. The apparatus was set up so that the researcher was outside subjects' field of vision, and assurances were given that they were not being judged on their singing ability; rather, they were encouraged to relax and follow their musical intuitions.

The materials were saved as MP3 files, and replayed using a Dell Dimension 3100C PC with a Lexicon Alpha soundcard and Harman Kardon speakers, which presented the stimuli to subjects at around 60dB. Responses were recorded through a Sony ECM-MS907 microphone connected to the same Dell PC. Vocal frequencies were measured in Hertz (Hz) using Praat, version

Figure 19. 'Practice' melody

Figure 20. The introductions played before the practice and stimulus melodies were heard

5.1.03 (Boersma and Weenink). This software was selected for its successful history in the field of psychological research (for example, Sergeant & Welch, 2009; Steinbeis, Koelsch, & Sloboda, 2006).

Procedure

In both the practice and experimental conditions, subjects were presented with the introduction, followed after a short pause by the first note of the melody. They were asked to sing the note that they thought would be most likely to come next. Then the first two notes of the melody were presented, and again, subjects were requested to sing the pitch that they thought would follow. This process continued until the penultimate note of the melody. After each response, subjects signified how confident they felt that their guess was correct on a Likert scale from 1 ('not at all confident') to 7 ('extremely confident'). They were encouraged to use the whole of the scale, but reserving endpoints for extreme cases. The experiment took about 20 minutes.

Initial data processing

The raw data comprised recordings of 1000 brief vocalizations (25 responses from each of 40 subjects). Each was measured in Hz, determined by taking the average frequency of the most consistent portion of the response. Occasionally there was significant variation: for example, where the vocal pattern started at a particular frequency, rose up to a higher one, then fell back again. Here, subjects' efforts were evaluated by an independent judge, who determined perceptually the point at which they seemed to settle on their intended pitch.

The frequencies obtained were assigned to categories from the D major scale, assuming equal temperament, and given that D4 (the D above middle C on the piano) = 294 Hz. Responses from male participants were transposed up an octave in music-notational terms, to facilitate male and female data being considered together.

Results and discussion

Combining the 40 subjects' responses (see Assumption 11), and scaling them so that the sum pertaining to each note is 1, yields the dataset shown in Figure 21.

The diversity is striking. The *number* of different predictions (N_P) varies from 3–12 (M = 7.36, SD = 1.96), and the *range* in scale degrees (R_P) from 5–13 (M = 8.68, SD 2.58). Combining these factors to give a 'coefficient of variability' (V_P), such that

$$V_P = \tfrac{1}{2} (N_P + R_P)$$

reveals that the specificity of expectations increases only by around a quarter, as Figure 22 shows. Despite designing the melody with a structure that was thought to be readily apprehensible (which was intended to make prediction easier as the music progressed), there is a high degree of variability in listeners' predictions. Even the last note, which is signalled both intraopusly (see Figure 11) and schematically (as the tonic at the end of a tonally and thematically symmetrical melody, whose first section concluded on the dominant), was not anticipated by three listeners.

The results show high inter-subject variability too (see Figure 23). With a potential maximum of 25, the number of correct predictions ranged from 0–19, M = 11.9, SD = 4.84.

Observed proportions of expectancies (rows) / **Note in stimulus melody** (columns)

	D4	E4	F♯4	A4	D4	E4	F♯4	A4	B4	A4	G4	B4	A4	D4	E4	F♯4	A4	D4	E4	F♯4	A4	E4	D4	C♯4	E4
B5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
A5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
G5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F♯5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
E5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D5	0.03	0.00	0.00	0.03	0.00	0.00	0.00	0.05	0.03	0.00	0.00	0.03	0.03	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00
C♯5	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
B4	0.00	0.00	0.03	0.23	0.00	0.00	0.00	0.13	0.13	0.05	0.03	0.13	0.05	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.75	0.00	0.00	0.00	0.00
A4	0.05	0.00	0.10	0.08	0.05	0.00	0.47	0.00	0.59	0.13	0.05	0.63	0.08	0.03	0.05	0.78	0.05	0.00	0.03	0.85	0.05	0.05	0.03	0.03	0.00
G4	0.03	0.00	0.62	0.25	0.03	0.03	0.11	0.05	0.08	0.26	0.00	0.05	0.31	0.03	0.03	0.10	0.00	0.05	0.00	0.03	0.13	0.13	0.00	0.00	0.00
F♯4	0.10	0.69	0.08	0.13	0.05	0.60	0.03	0.03	0.03	0.41	0.68	0.05	0.05	0.03	0.75	0.03	0.00	0.00	0.80	0.08	0.00	0.33	0.00	0.05	0.00
E4	0.48	0.10	0.00	0.05	0.36	0.08	0.11	0.03	0.03	0.03	0.15	0.03	0.00	0.58	0.05	0.05	0.00	0.78	0.05	0.03	0.00	0.10	0.15	0.33	0.00
D4	0.15	0.03	0.15	0.13	0.18	0.20	0.26	0.53	0.03	0.05	0.08	0.05	0.33	0.25	0.05	0.03	0.80	0.10	0.08	0.00	0.03	0.35	0.20	0.33	0.93
C♯4	0.00	0.05	0.00	0.00	0.18	0.00	0.03	0.05	0.00	0.05	0.00	0.00	0.05	0.05	0.03	0.00	0.03	0.00	0.00	0.00	0.00	0.03	0.33	0.10	0.05
B3	0.08	0.05	0.00	0.00	0.03	0.03	0.00	0.03	0.03	0.00	0.03	0.03	0.03	0.00	0.00	0.00	0.03	0.03	0.03	0.03	0.00	0.03	0.20	0.13	0.00
A3	0.05	0.05	0.03	0.10	0.08	0.08	0.00	0.05	0.00	0.00	0.00	0.00	0.05	0.03	0.03	0.03	0.05	0.03	0.03	0.00	0.00	0.00	0.03	0.03	0.00
G3	0.03	0.03	0.00	0.00	0.03	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.03	0.00	0.03	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.03	0.00
F♯3	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.03	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.03
E3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00
D3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C♯3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Np	**9**	**7**	**6**	**9**	**10**	**6**	**6**	**12**	**9**	**8**	**6**	**9**	**10**	**8**	**8**	**6**	**7**	**6**	**6**	**5**	**6**	**7**	**8**	**8**	**3**
Rp	**13**	**6**	**8**	**10**	**9**	**6**	**5**	**13**	**9**	**10**	**7**	**13**	**11**	**10**	**8**	**7**	**11**	**7**	**7**	**6**	**13**	**6**	**9**	**8**	**5**
Vp	**11**	**6.5**	**7**	**9.5**	**9.5**	**6**	**5.5**	**12.5**	**9**	**9**	**6.5**	**11**	**10.5**	**9**	**8**	**6.5**	**9**	**6.5**	**6.5**	**5.5**	**9.5**	**6.5**	**8.5**	**8**	**4**

(shaded) = pitch that actually occurred

Figure 21. Observed, scaled responses

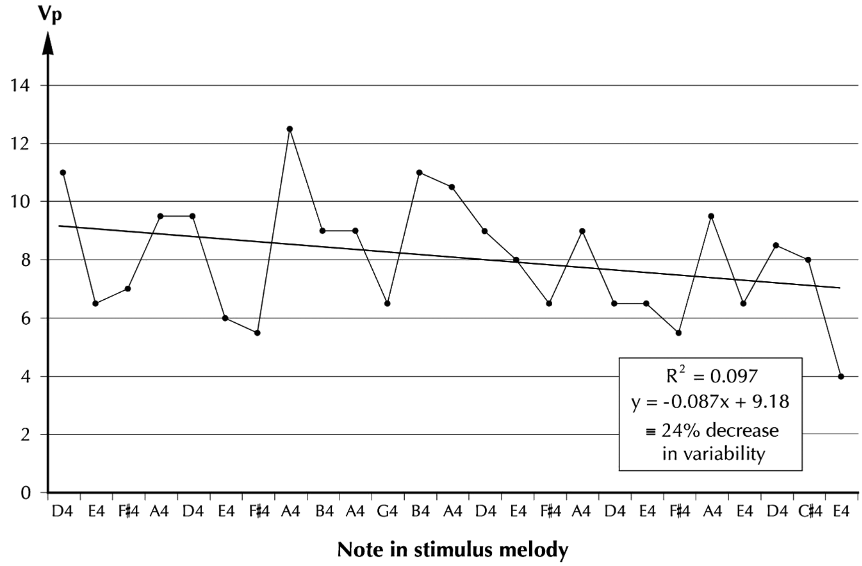


Figure 22. Decrease in the variability of responses over the course of the melody

There are no significant effects for age, years of musical training, or the time reportedly spent listening to music.

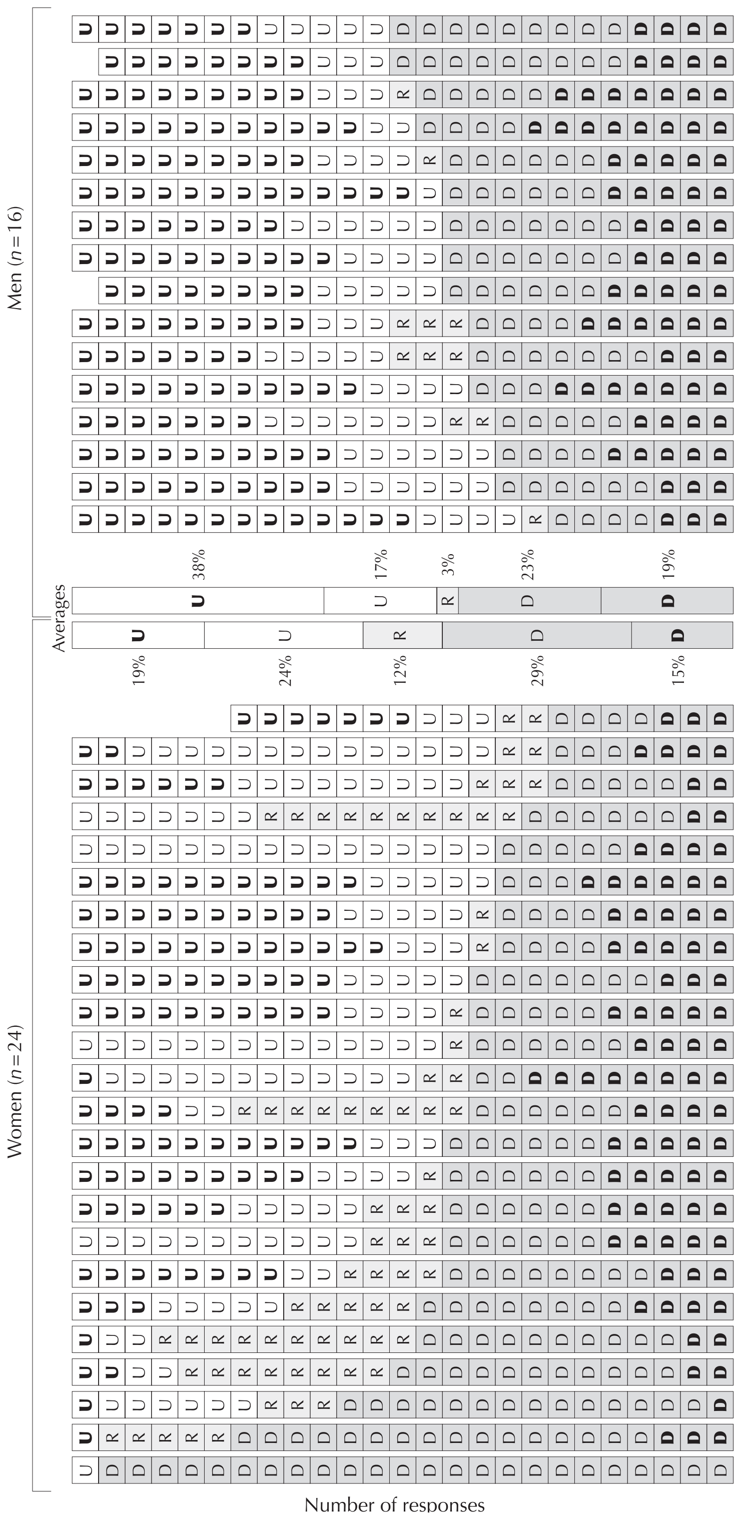
However, gender does appear to be important, with a significant difference between men's (n = 16, M = 14.3, SD = 2.75) and women's (n = 24, M = 10.3, SD = 5.28) success in anticipating what came next: $t(38) = 2.83$, $p = .0075$. As the groups of male and female subjects did not differ significantly in terms of age, musical training, or time spent listening to music, one has to look for other explanations. Clearly, the relatively small numbers of men and women involved may have resulted in unrepresentative samples – an issue that only larger-scale replication could resolve. Then, there was an experimental difference between the two groups, whereby the octave in which responses were elicited was different for men and women (Constraint 7; see Figure 11). Were tessitura a significant factor in the formation of expectations, there would be a difference in the frequencies with which intervals of different polarities were predicted (up, down, or neutral), according to where the stimulus was pitched, since, *ceteris paribus*, the lower a given note, the greater the probability that the one following will be higher, and *vice versa* (Huron, 2006, pp. 80ff). Following this reasoning, women's predictions (whose stimulus was higher) should contain a greater proportion of descending intervals, while men's should be more likely to ascend. Analysis by intervallic polarity does indeed show a difference in the patterns of response by gender, but these are not straightforward (see Figure 24).

With regard to descending intervals, the proportions predicted by women (M = 11, SD = 3.8) and men (M = 10.6, SD = 1.6) are very similar: 44% and 42% respectively, which suggests no effect of tessitura. The data pertaining to ascending intervals paint a different picture, though: women's predictions (M = 10.7, SD = 4.88; 43%) and men's (M = 13.6, SD = 1.71; 55%) differ significantly, $t(38) = 2.28$, $p = .029$, supporting the tessitura hypothesis. Moreover, as the proportion of ascending intervals to descending is 3:2, this could potentially explain men's greater success in prediction.

'Actual' note in melody
versus
predicted note in melody

Participant	E4	F♯4	A4	D4	E4	F♯4	A4	B4	A4	G4	B4	A4	D4	E4	F♯4	A4	D4	E4	F♯4	A4	E4	D4	C♯4	E4	D4	Number of correct predictions	Sex	Age	Years of musical traning	Minutes per day spent listening to music
1	A3	F♯4	A4	F♯4	E4	F♯4	A4	D4	A4	F♯4	F♯4	B4	D4	E4	F♯4	A4	D4	E4	F♯4	A4	B4	D4	E4	D4	D4	16	F	21	–	>60
2	F♯4	G3	G4	G4	G3	E4	G4	G3	A4	C♯4	A4	B4	A3	C♯4	G3	G4	A3	A3	D4	A4	B4	B3	A3	G3	C♯4	2	F	26	–	30-60
3	F♯4	F♯4	G4	G4	F♯4	F♯4	A4	D4	A4	G4	F♯4	A4	G4	E4	F♯4	A4	D4	E4	F♯4	A4	G4	G4	C♯4	F♯4	D4	15	F	22	>8	30-60
4	E4	D4	G4	F♯4	E4	F♯4	C♯4	C♯4	A4	G4	F♯4	F♯4	A4	E4	F♯4	A4	D4	E4	F♯4	A4	B4	D4	C♯4	C♯4	D4	15	F	22	<2	>60
5	E4	F♯4	G4	G4	E4	F♯4	A4	D4	A4	A4	F♯4	A4	D4	E4	F♯4	A4	D4	E4	F♯4	A4	B4	G4	B3	E4	D4	17	F	24	>8	>60
6	F♯4	F♯4	G4	B4	F♯4	F♯4	D4	D4	A4	D4	F♯4	A4	G4	E4	F♯4	A4	D4	E4	F♯4	A4	B4	F♯4	E4	E4	D4	13	F	24	>8	30-60
7	E4	F♯4	G4	G4	A3	A3	—	D4	G4	G4	F♯4	A4	B3	E4	F♯4	A3	D4	E4	F♯4	A4	B4	F♯4	B3	E4	D4	12	M	36	5-8	30-60
8	E4	F♯4	G4	B4	C♯4	D4	A4	D4	C♯5	F♯4	F♯4	A4	B4	E4	F♯4	G4	D4	G4	F♯4	E4	G4	F♯4	C♯4	E4	D4	11	F	27	5-8	>60
9	D4	—	—	G4	E4	F♯4	—	D4	—	—	F♯4	E4	—	D4	F♯4	A4	D4	E4	F♯4	A4	B4	D4	D4	D4	D4	10	F	24	<2	<30
10	F♯4	F♯4	G4	B4	E4	D4	D4	D4	A4	A4	F♯4	A4	D5	D4	F♯4	A4	D4	E4	F♯4	A4	B4	F♯4	B3	D4	D4	11	M	50	–	<30
11	A4	F♯4	G4	B4	A3	F♯4	A4	B4	D5	F♯4	F♯4	A4	F♯4	E4	F♯4	A4	D4	E4	F♯4	A4	B4	F♯4	B3	A3	D4	13	M	25	>8	>60
12	B3	E4	G4	E4	D4	D4	D4	E4	B4	B4	E4	A4	A4	D4	D4	A4	D4	E4	D4	A4	B4	D4	D4	D4	D4	7	F	31	–	30-60
13	E4	F♯4	D4	F♯4	E4	F♯4	D4	D5	A4	G4	F♯4	D4	G4	A4	F♯4	D4	D4	E4	F♯4	A4	B4	F♯4	F♯4	E4	D4	12	M	55	2-4	>60
14	B3	B3	F♯4	A4	B3	F♯4	A4	B3	A4	A4	E4	A4	G4	D4	D4	G4	A3	D4	B3	G4	A4	E4	B3	C♯4	C♯4	4	F	23	–	>60
15	B3	E4	D4	A3	A3	B3	E4	F♯3	B3	F♯3	B3	B3	A3	A3	C♯4	E4	E3	G3	A3	B3	E3	C♯4	B3	B3	F♯3	0	F	30	–	>60
16	D4	F♯4	G4	G4	D4	F♯4	A4	D4	A4	G4	E4	A4	D4	E4	F♯4	A4	D4	E4	F♯4	A4	B4	E4	B3	D4	D4	15	M	60	–	>60
17	E4	F♯4	G4	D4	E4	F♯4	D4	D4	C♯5	F♯4	F♯4	A4	D4	E4	F♯4	A4	D4	E4	F♯4	A4	B4	A4	E4	F♯4	D4	15	M	27	–	30-60
18	E4	F♯4	G4	D4	E4	F♯4	A4	D4	A4	F♯4	F♯4	A4	D4	E4	F♯4	A4	D4	E4	F♯4	A4	B4	D4	C♯4	B3	D4	19	M	30	2-4	30-60
19	E4	F♯4	G4	D5	E4	F♯4	E4	D4	A4	F♯4	E4	A4	D4	E4	F♯4	A4	D4	E4	F♯4	A4	G4	G4	F♯4	C♯4	D4	15	F	27	5-8	30-60
20	E4	F♯4	D4	B4	E4	F♯4	A4	D4	A4	G4	F♯4	B4	D4	D4	F♯4	A4	D4	E4	F♯4	A4	B4	D4	C♯4	D4	D4	17	F	24	–	>60
21	E4	F♯4	G4	G4	C♯4	F♯4	A4	G4	G4	G4	F♯4	A4	G4	E4	F♯4	A4	D4	E4	F♯4	A4	B4	F♯4	C♯4	E4	D4	16	F	24	>8	30-60
22	D4	F♯4	F♯4	A4	D4	D4	E4	D4	B4	F♯4	B4	F♯4	G4	D4	F♯4	A4	D4	D4	E4	A4	B4	E4	D4	A4	D4	7	F	23	5-8	<30
23	G3	F♯4	D4	C♯5	D4	F♯4	D4	D4	B4	F♯4	D4	G4	C♯4	D4	F♯4	A4	D4	E4	F♯4	A4	B4	D4	D4	E4	D4	11	F	29	2-4	<30
24	E4	F♯4	G4	D4	E4	F♯4	D4	D4	A4	G4	F♯4	A4	D4	E4	F♯4	A4	D4	E4	F♯4	A4	B4	F♯4	E4	E4	D4	18	M	28	<2	>60
25	E4	F♯4	A4	G4	C♯4	F♯4	A4	B4	A4	F♯4	F♯4	A4	F♯4	E4	F♯4	A4	D4	E4	F♯4	A4	G4	F♯4	C♯4	E4	D4	18	M	40	>8	>60
26	F♯4	C♯4	D4	B4	C♯4	F♯4	E4	B4	B4	F♯4	F♯4	A4	B4	E4	F♯4	A4	D4	E4	F♯4	A4	B4	D4	C♯4	E4	D4	14	M	58	5-8	30-60
27	E4	F♯4	G4	B4	C♯4	F♯4	A4	D4	A4	A4	F♯4	A4	D4	E4	F♯4	A4	D4	E4	F♯4	A4	B4	D4	B3	D4	D4	16	M	22	5-8	<30
28	E4	F♯4	G4	B4	E4	F♯4	A4	B4	A4	G4	F♯4	B4	G4	E4	F♯4	A4	D4	E4	F♯4	A4	B4	F♯4	E4	B3	D4	16	M	59	5-8	30-60
29	E4	F♯4	G4	A3	—	A3	G4	A3	A4	F♯4	F♯4	A4	G4	E4	A4	A4	D4	E4	F♯4	A4	B4	F♯4	C♯4	E4	D4	13	M	32	2-4	30-60
30	E4	C♯4	A3	D4	F♯3	D4	F♯4	E3	E4	D4	D4	E3	G3	C♯4	E4	E4	B3	B3	D4	F♯4	A4	D4	F♯3	C♯4	D4	4	F	33	<2	30-60
31	D4	E4	D4	A4	D4	E4	A4	D4	A4	E4	F♯4	G4	D4	D4	F♯4	A4	A4	D4	A4	A4	D4	D4	D4	D4	D4	8	F	27	–	30-60
32	G4	F♯4	A4	D4	A4	D4	D4	D4	A4	F♯4	D4	B4	D4	E4	F♯4	A4	D4	E4	F♯4	A4	B4	D4	C♯4	E4	D4	16	M	30	2-4	30-60
33	A4	A3	G4	E4	A4	D4	A4	B4	D4	A4	F♯4	D5	A4	D5	A4	A4	D4	E4	F♯4	A4	D5	A4	D4	D4	D4	8	F	73	–	30-60
34	A3	E4	F♯4	A3	D4	A3	A4	A3	F♯4	C♯4	F♯4	D4	C♯4	D4	E4	F♯4	A4	D4	E4	F♯4	G4	D4	D4	D4	D4	3	F	21	2-4	30-60
35	E4	F♯4	G4	B4	C♯4	F♯4	A4	D4	A4	G4	F♯4	A4	D4	E4	F♯4	A4	C♯4	E4	F♯4	A4	B4	G4	C♯4	E4	D4	17	F	25	>8	>60
36	D4	F♯4	A4	F♯4	D4	E4	G4	G4	A4	F♯4	F♯4	A4	G4	F♯4	G4	A4	D4	E4	F♯4	A4	B4	E4	C♯4	B3	D4	11	F	76	<2	30-60
37	E4	F♯4	G4	G4	E4	D4	D4	D4	B4	F♯4	E4	A4	G4	D4	F♯4	A4	D4	E4	F♯4	A4	B4	F♯4	D4	D4	D4	10	M	27	>8	>60
38	D5	B3	B4	A3	C♯4	F♯4	A4	C♯4	C♯5	F♯4	F♯4	A4	G4	E4	A3	A4	D4	E4	F♯4	A4	B4	F♯4	C♯4	B3	D4	11	M	37	>8	>60
39	D4	A3	G4	G4	G4	G4	G4	D5	A4	B4	A4	A4	D4	G4	F♯4	G4	B4	G4	F♯4	F♯4	B4	D4	A4	D4	D4	6	F	22	5-8	<30
40	E4	F♯4	G4	F♯4	E4	F♯4	D4	F♯4	G4	F♯4	E4	A4	G4	E4	F♯4	A4	D4	E4	F♯4	A4	B4	G4	E4	D4	D4	13	F	65	5-8	30-60

Figure 23. Individual responses



Sums: $n = 594$, D = 264 (**D** = 90); R = 74; U = 256 (**U** = 114)
Averages: D = 11.0 (**D** = 3.8); R = 3.2; U = 10.7 (**U** = 4.8)

Sums: $n = 398$, D = 169 (**D** = 77); R = 11; U = 218 (**U** = 150)
Averages: D = 10.6 (**D** = 4.8); R = 0.7; U = 13.6 (**U** = 9.4)

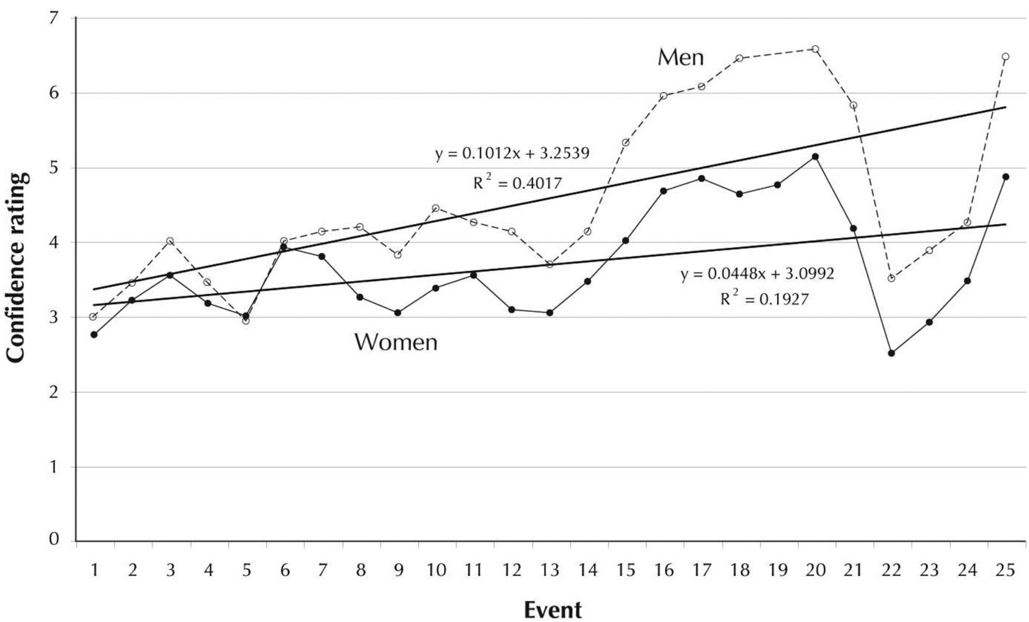
Key: D = downward interval (incorrect); **D** = downward interval (correct); R = repetition of same note (invariably incorrect); U = upward interval (incorrect); **R** = upward interval (correct).

Figure 24. Intervallic polarity of responses

There are two confounding factors, however. First, men's *proportion* of successful ascending predictions (150 out of 218, or 69%) was significantly greater than women's (114 out of 256, or 44%), $\chi^2(1, n = 474) = 28.12$, $p < .0001$. (Men's proportion of successful *descending* predictions, being 77 out of 169, or 45%, was also greater, at 90 out of 264, or 34%, although here the difference was less marked, $\chi^2(1, n = 453) = 8.76$, $p = .003$.) Second, the different proportion of ascending predictions can largely be accounted for by women's tendency to expect pitches to be repeated (although none in fact was) ($M = 3.1$, $SD = 3.20$; 12%), which was significantly greater than that for men ($M = 0.7$, $SD = 1.08$; 3%): $t(38) = 2.88$, $p = .007$.

The *confidence* with which men and women made their predictions differed significantly too. On a scale of 1 (low) to 7 (high), women's confidence ratings were $M = 3.68$, $SD = 1.42$, while men's were both higher and more consistent, with $M = 4.57$, $SD = 0.96$: $t(38) = 2.19$, $p = .035$. Although men were almost invariably more confident than women (after only one event, note 5, were men less so, and then by just 0.07 of a point), and while the confidence of both sexes grew through the tests, men increased in confidence almost twice as much as women: by around two points on the Likert scale as opposed to one (see Figure 25). It could be that, as men's early predictions proved to be correct more often than women's, their confidence grew more strongly.

In summary, while tessitura may explain some of the differences in men's and women's expectancy profiles, it is not the only factor – and it could be that the results reflect a difference in the way that males and females process and predict musical structure. This possibility has some neuropsychological support: Koelsch, Maess, Grossmann, and Friederici (2003) found



women	2.75	3.21	3.54	3.17	3.00	3.92	3.79	3.25	3.04	3.38	3.54	3.08	3.04	3.46	4.00	4.67	4.83	4.63	4.75	5.13	4.17	2.50	2.92	3.46	4.83
men	3.00	3.44	4.00	3.44	2.93	4.00	4.13	4.19	3.81	4.44	4.25	4.13	3.69	4.13	5.31	5.94	6.06	6.44	6.50	6.56	5.81	3.50	3.88	4.25	6.44
difference	**0.25**	**0.23**	**0.46**	**0.27**	**-0.07**	**0.08**	**0.33**	**0.94**	**0.77**	**1.06**	**0.71**	**1.04**	**0.65**	**0.67**	**1.31**	**1.27**	**1.23**	**1.81**	**1.75**	**1.44**	**1.65**	**1.00**	**0.96**	**0.79**	**1.60**

Figure 25. The changes in confidence ratings of predictions for men and women

that an electrophysiological correlate of music-syntactic processing is generated bilaterally in females, with right hemispheric predominance in males. There appear to be attitudinal differences too, and the surety with which predictions are made may have impacted on expectational 'success'. For example, women's lower levels of predictive confidence may have resulted in the greater frequency with which they anticipated repetition (which never actually occurred) rather than change. It is possible that such differences may have arisen from opposing personality traits, stereotyped as 'male' and 'female' in Western culture. These are intriguing areas for future research.

The range and variability of data across all subjects (particularly in the case of women, whose patterns of expectation were more varied, as the differences in standard deviation show) raise two phenomenological issues. First, the variation in responses could be held to argue against a common mechanism of expectation. Second, as over half the predictions (52%) were incorrect, could expectation of this kind be part of the 'typical' listening experience (Ockelford, 2006, p. 135), or is it a feature of musical metacognition that was induced experimentally?

The zygonic model (versions **Z1** and **Z2**) can potentially accommodate both these concerns. First (see Research Question 1), it predicts that materials pertaining to 'current structures, within groups' and 'previous structures, schematically encoded' interact to produce a *general* sense of what may follow. Hence, one would *expect* listeners' responses to differ, and would anticipate that, taken together, they would reflect the probabilistic nature of the model. Moreover, if we hypothesize that listeners' expectations usually occur nonconsciously (notwithstanding our capacity to draw these into conscious awareness), then the second, aesthetic, issue can potentially be resolved (although this would require further empirical work). Therefore, the key question is how well **Z1** and **Z2** fit the observed data. The predictions of all three models and subjects' responses are summarized in Figure 26.

Model **Z1** produces broader spans of values than those of subjects' expectations, though since, at the extremes, the figures are small, the impact on similarity is modest, and the mean difference of predicted and observed values is only 0.065 ($SD = 0.022$). With **Z2**, the mean difference of predicted and observed values is 0.059 ($SD = 0.015$). The difference is significant: $t(24) = 2.58$, $p = .017$. This suggests that acknowledging the potential influence of the four most recent pitches (where they exist) provides a more accurate model of listeners' expectations than the effect of the most recent value alone. And it is evidence to support the hypothesis postulated in Research Question 1 that, *ceteris paribus*, expectations arising from 'current structures, within groups' (the four preceding notes) and 'previous structures, schematically encoded' (the relative frequencies of past intervallic occurrence; see Figures 13 and 14 earlier) interact to produce a *general* sense of what may follow.

Research Question 2 asks whether there is support for the hypothesis that expectations arising from 'current structures, between-groups' produce a *specific* sense of what is to come. Evidence can be sought by comparing the coefficients of variability (V_P) of those pitches that the zygonic analysis shown in Figure 11 indicates could be predicted through between-group relationships and those that could not, and the data set out in Figure 27 indicate that there is indeed a significant difference in the average V_P values pertaining to each category, $t(23) = 2.59$, $p = .016$. This finding is supported by the differences in confidence with which subjects reported their predictions were made: where between-group structures were available, anticipation was significantly more assured than where they were not, $t(23) = 3.21$, $p = .004$.

Research Question 3 asked whether there was evidence to support the notion that expectations arising from (a) 'current structures, within groups' and 'previous structures, schematically encoded' interact with (b) 'current structures, between-groups', such that (b) lends

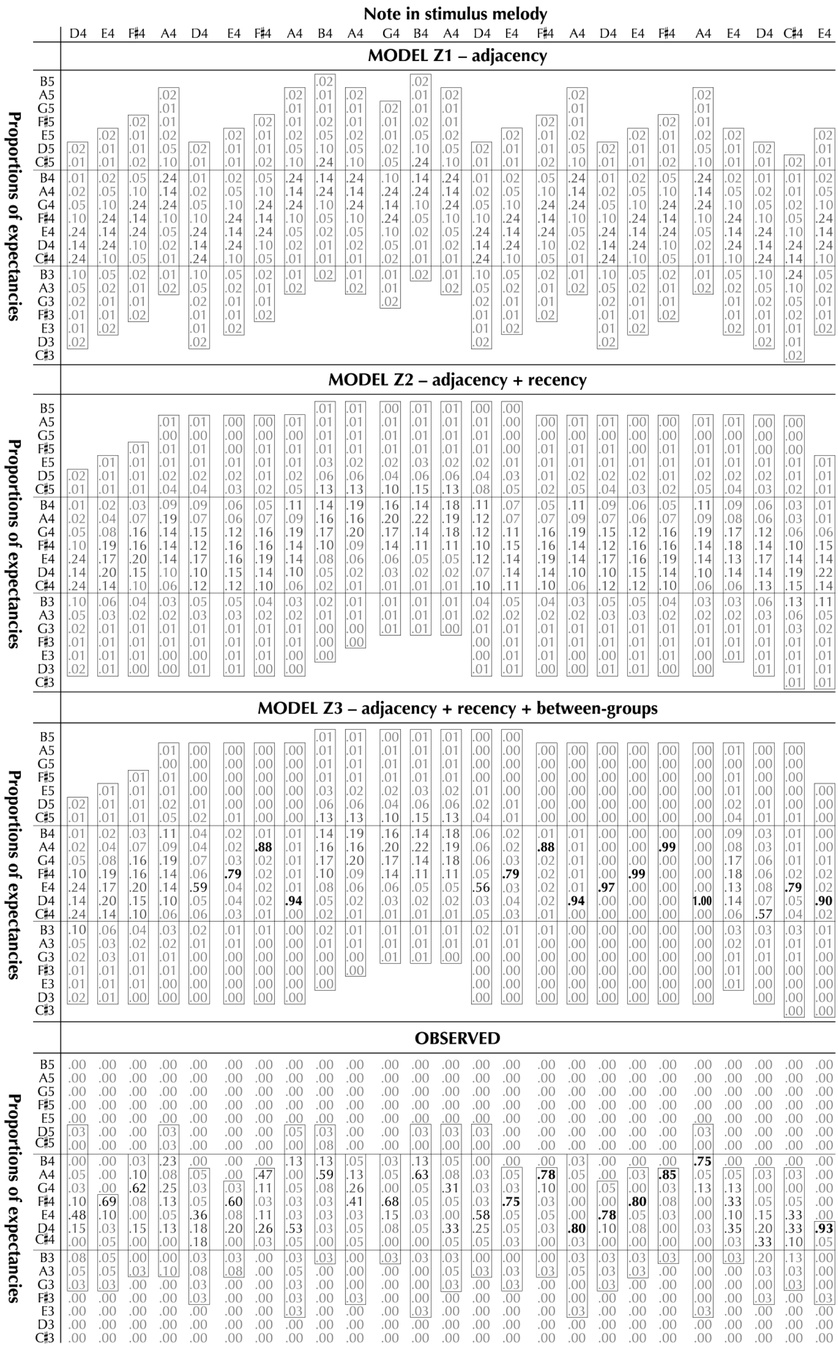


Figure 26. Predicted and observed responses

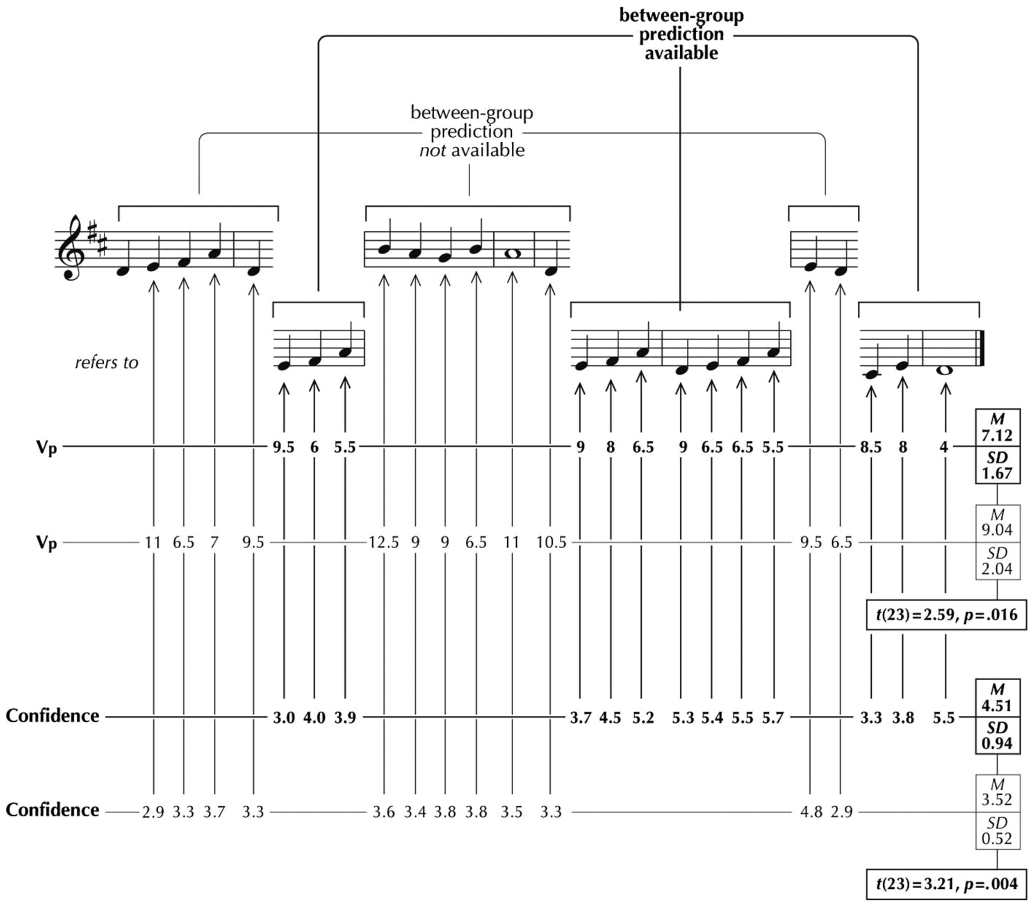


Figure 27. Variability and confidence of prediction differ significantly with the existence or non-existence of between-group structure

greater specificity to (a), and (a) grounds (b) in a local context. Model **Z3** was constructed to test this hypothesis, fusing all three anticipatory elements; therefore, the degree to which the predictions of **Z3** and the observed responses correspond will reflect the extent to which the theory is supported.

Analysis of the data shown in Figure 26 indicates that the average difference between predicted and observed responses was only $M = 0.039$, $SD = 0.019$, and that this differed significantly from **Z2**: $t(24) = 3.79$, $p = .0009$. Hence the data support aspects **A1**, **A3**, and **B2** of the zygonic model. Figure 28 provides a compelling visual metaphor of the changing nature of people's expectations as they responded to the melody over time, and compares them with the outputs of **Z3**.

Taking grand averages of these results reveals a high degree of correlation between the model and subjects' responses (Figure 29).

There are differences, however, which indicate limitations of the model and suggest potential modifications, as well as suggesting possible improvements in the experimental design. In

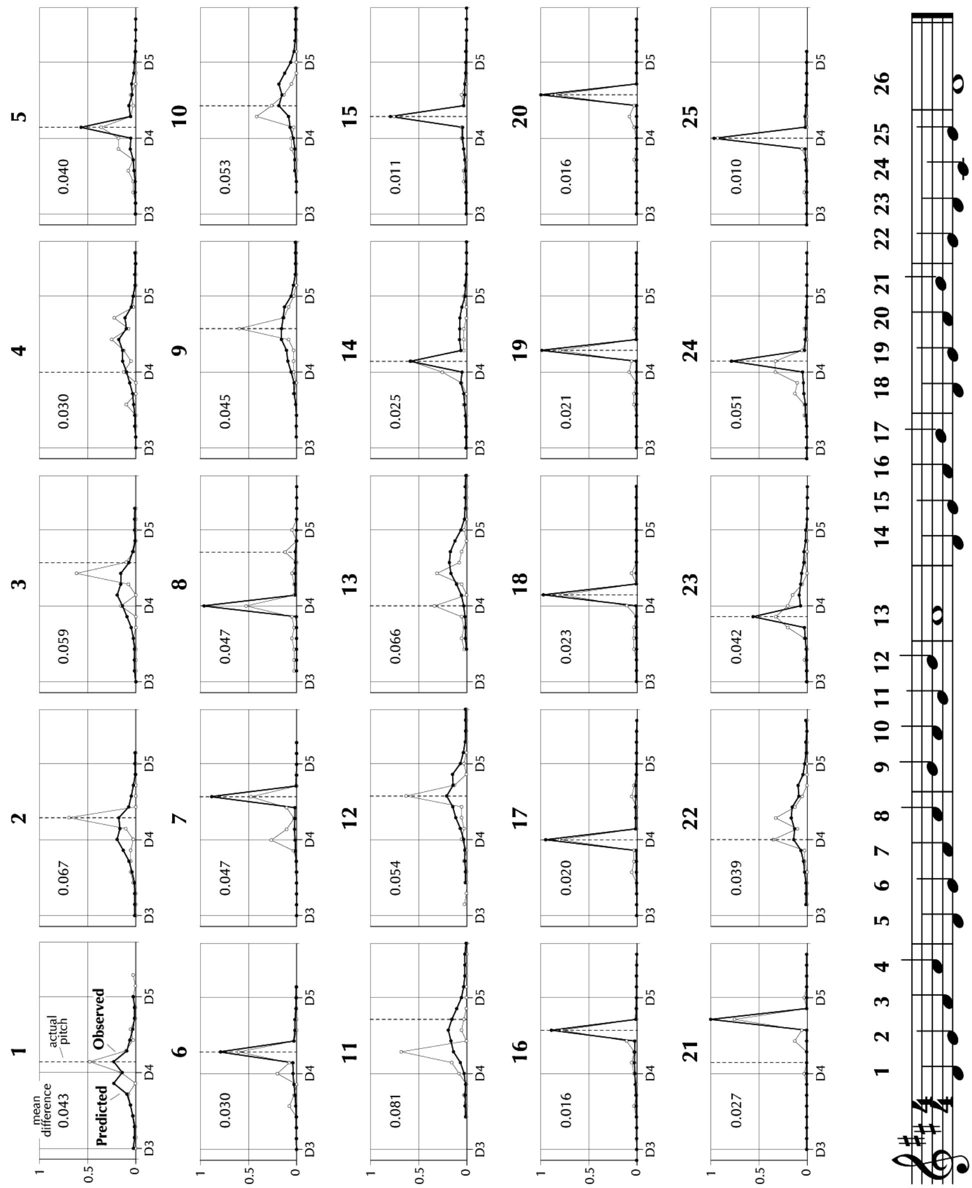


Figure 28. Correlations between Z3 outputs and observed responses, event by event

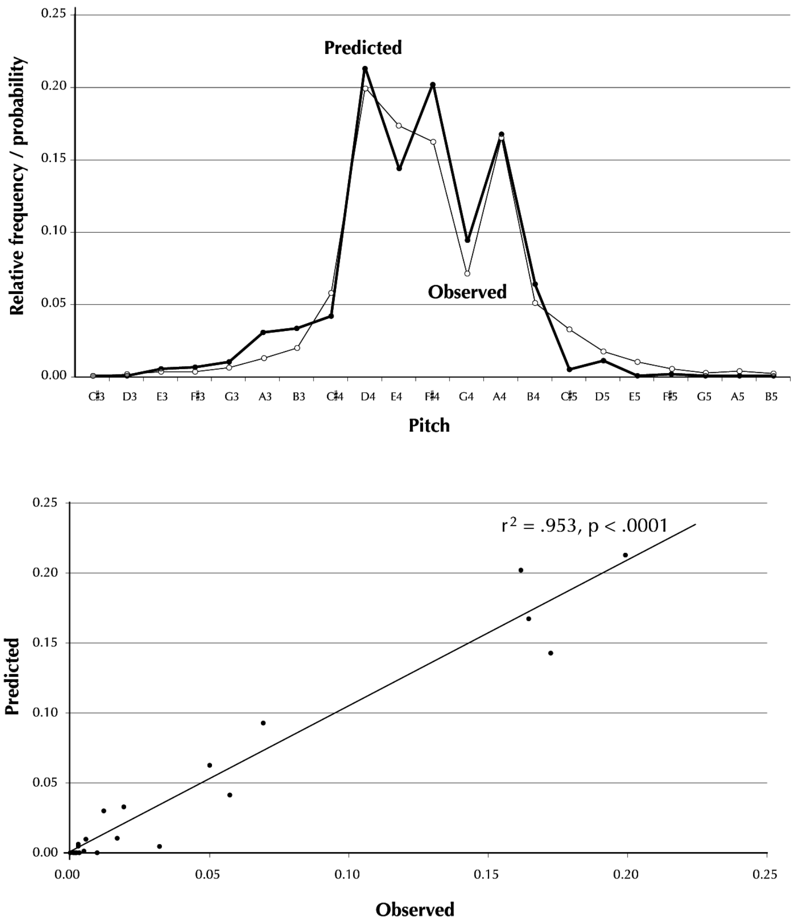


Figure 29. Correlation between the means of all Z3 outputs and observed responses, note by note

relation to the first three events, **Z3** predicts a broad spread of values centred on the pitches that are heard – a consequence of the primary zygonic relationships upon which it is based. While subjects also anticipated a range of values, there was a bias towards expecting a pitch one scale degree higher than the one heard. Why? With regard to the first event, two factors may have played a part. First, some listeners may have internalized the tendency of melodies to begin by ascending (Huron, 1996, 2006, pp. 86ff), creating a schema that they intuitively brought to bear. This could feature in future models as an additional factor in **B2**. And it may be that this tendency was reinforced by the introductory material, which began with a rising scale. Although the subsequent descent was intended to counter any such effect, the introductory gesture may have reinforced the sense of an 'arch' contour typical of many melodies (Huron, 2006). Future experiments could use introductory materials that gave an unambiguous sense of key without incurring melodic movement.

The second stimulus pitch (a scale step higher than the first) would have supported the expectations of the 49% of subjects who anticipated the ascent, and suggested to all listeners

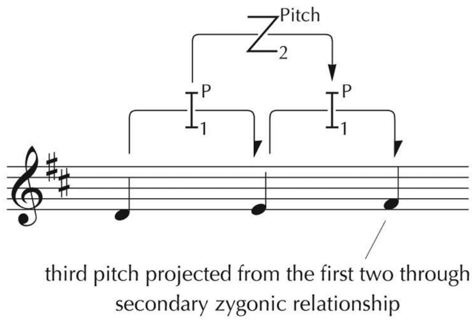


Figure 30. Anticipation through secondary zygonic structure not factored in to **Z3**

that the melodic opening was based on stepwise upward movement. There were three possible sources of such a projection: the schema of initial melodic ascent, the introductory material, and, most immediately, the opening interval – each reliant on secondary zygonic relationships (see Figure 30).

However, in relation to **A1** ('current structures encoded in working memory'), secondary relationships were not tested here (see Constraint 8), as there was no evidence as to how primary and secondary relationships arising from current structures may interact. Given the current findings, however, it may be possible to include this element in future empirical work. The fact that listeners recognized the pattern of ascent is shown by their tendency (62%) to anticipate G4 as the fourth event. However, the A4 that actually occurred broke the pattern, and at this stage subjects projected a range of values, whose profile was close to that predicted by **Z3**.

From Event 5 (see Figure 28), evidence emerges of between-group expectation, reflected in the peak shared by subjects' responses and **Z3**, although the effect in the zygonic model is stronger (58% as opposed to 36%), with a consequent underestimate of the probabilities assigned to other pitches. That is, in the zygonic model, the balance between projections based on adjacency and recency, and between-group relationships, moves in too pronounced a fashion towards the latter. This is affirmed in differences between the subjects' and **Z3**'s projections that follow Events 6 and 7. In future models, the ratio between the competing forms of expectation could be adjusted (cf. Figure 18).

After the second A4 (Event 8), both the subjects and **Z3** predicted a continuation of the pattern (a further D4), though the melody took a different course, rising to B4. This caused a collapse at Event 9 of specificity in the model and, to a certain extent, in the responses, although most subjects anticipated A4, providing further evidence of secondary zygonic relationships being used in anticipation – here through inversion.

After Event 10, with no between-group relationships in play, the model predicts a broad range of values – partly matched by the profile of subjects' responses, although a clear majority (41%) anticipate F♯4. It is unclear why this prediction dominates, since the secondary zygonic structural logic in evidence up to this point would have projected G4 (although this was the second most common expectation at 26%).

At Event 11, secondary zygonic structure reasserts itself, and the majority of subjects (68%) anticipate F♯4 as a continuation of the descent from B4, A4, and G4 (Figure 31), while **Z3** offers a general profile of expectation. At Event 12, A4 is accurately anticipated

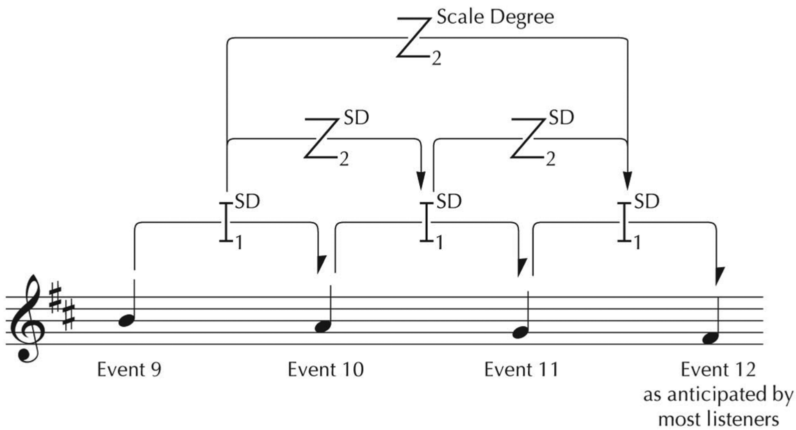


Figure 31. Further anticipation through secondary zygonic structure not factored into **Z3**

by most listeners (63%), explicable through a combination of primary and secondary zygonic projection from notes 10 and 11. **Z3** also predicts A4 as the most likely successor, though here the effect is more muted, perhaps because only primary relationships are taken into account.

At Event 13, **Z3** projects a spread of values centred on A4, whereas around a third of subjects anticipated a continuation of the descent from B4 and A4 to G4 (implying further secondary zygonic organization) and a third expected a return to D4, presumably a schematic response to the implied half cadence (ending on the dominant), which often presages a return to the tonic. This potential characteristic of **B2** could be captured in future modelling.

At Event 14, between-group structures feature again, whose effects on expectation gradually gain in strength among listeners and **Z3**: the longer the repetition of the opening phrase continues, the surer its continuation is felt to be, although the zygonic model again predicts individual values more strongly than the subjects, suggesting that some modification to the adjacency-recency/between-group ratio is required.

At Event 21, both listeners and **Z3** project B3, but the melody deviates from its previous course, and the relative certainties of between-group anticipation again dissipate into generality at Event 22. At Event 23, the transposition of the motif first heard in bar 3 is acknowledged by subjects and the zygonic model alike, a recognition that grows more assured at Event 24, though, once more, **Z3** is firmer in its between-group prediction. At Event 25 (the penultimate note), there is the strongest correlation of all as **A1**, **B2**, and **C3** act together to predict an unambiguous return to the tonic.

In summary, **Z3** is broadly effective in modelling subjects' melodic expectations in the context of the melody presented here, though the results indicate three areas where refinements could be made. First, the ratio between **A1** and **A3** should be adjusted, so that the impact of between-group relationships is less pronounced, particularly at the beginning of groups. Second, the potential effect of within-group secondary zygonic relationships (**A1**) needs to be included. Third, the effects of schemata pertaining to melodic contour and symmetry (**B2**) should be recognized. How these features interact with the adjacency and recency in relation to primary relationships, and links between groups, future empirical work could determine.

Conclusion

This article describes preliminary research that tested aspects of the zygonic model of expectation in music. Broad support was found for the theory, particularly the notion that, at a first hearing, *general* expectations may arise in response to within-group and schematically encoded structures, and *specific* expectations may be stimulated by the repetition or transposition of groups of notes. The proposition that specific projections are enabled by veridical memories was not tested. Suggestions were made for refining the model, such as incorporating more within-group and schematic information. The study had limitations, not least the fact that only 40 listeners' responses were tested in relation to one melody, whose design will inevitably have influenced the results. The findings confirm that humans *can* project what is coming next when listening to music, but provides no evidence that they usually *do*. Indeed, the fact that most projections made in the absence of inter-group information subsequently proved to be incorrect throws into doubt whether, beyond having a general sense of what is coming next, listeners actively seek to anticipate the future in first-time hearings, since this would presumably have a negative aesthetic impact.

An unexpected finding was that men and women appear to predict (and therefore process) musical structure differently, and future research could aim to establish to what extent commonalities and differences in musical expectation may exist between distinct sub-populations. Such work may take us a step closer to the venerable and tantalizing issue of the extent to which we all experience a piece of music in the same way or idiosyncratically.

Acknowledgments

With thanks to the volunteers who took part in the research, to Dr. Evangelos Himonides, Institute of Education, University of London, for his comments on a draft of the manuscript, and to the two anonymous reviewers for their telling insights into earlier versions of this article.

Notes

1. Zygonic relationships such as those depicted in Figure 1 offer, at best, a highly simplified version of certain cognitive events that are hypothesized to take place during meaningful participation in musical activity. Moreover, the single concept of a zygon bequeaths a vast perceptual legacy, with many manifestations: potentially involving any perceived aspect of sound; existing over different periods of perceived time; and operating within the same and between different pieces, performances, and hearings. Zygons may function in a number of different ways: *reactively*, for example, in assessing the relationship between two extant qualities of sounds; or *proactively*, in ideating an attribute as an orderly consequence of one that has been heard (the notion that lies at the heart of expectation as it is held to function in the current article). Zygons may operate between anticipated or remembered sounds, or even those that are wholly imagined, only ever existing in the mind. Hence there is no suggestion that the one concept is cognitively equivalent in all these manifestations, but that it is *logically* so.
2. Desmond Sergeant, private communication.
3. Observe that, here, the repetition of a pitch – 5th octave B – would not be expected to signal a group boundary, since the two notes occupy different positions in the bar, implying a conflict with the prevailing metre. While this is perfectly possible (as in the chorus of Gershwin's aptly-named *Fascinating Rhythm*, for example) the ear would need a signal, such as a break in the sound or a dynamic accent, to detect the syncopation prospectively (rather than in retrospect).
4. Observe that, in this first iteration of the model, the adjacency/recency quantification takes into account only *pitch* relationships, and ignores other factors that are likely to exert a perceptual influence, such as 'relative metrical location' or 'RML' (the position of the note in the bar; see Ockelford, 2009, p. 71). Future models could be refined to acknowledge cross-domain effects of this type.

References

Barlow, H., & Morgenstern, S. (1948). *A dictionary of musical themes*. London: Faber & Faber.

Bever, T. (1988). A cognitive theory of emotion and aesthetics in music. *Psychomusicology*, *7*, 165–175.

Bharucha, J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception*, *5*, 1–30.

Bharucha, J. (1994). Tonality and expectation. In R. Aiello with J. Sloboda (Eds.), *Musical perceptions* (pp. 213–239). New York: Oxford University Press.

Bharucha, J. (1999). Neural nets, temporal composites, and tonality. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 413–440). New York: Academic Press.

Boulez, P. (1971). *Boulez on music today* (S. Bradshaw & R. R. Bennett, Eds.). London: Faber & Faber. (Original work published in 1963)

Bregman, A. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.

Carlsen, J., Divenyi, P., & Taylor, J. (1970). A preliminary study of perceptual expectancy in melodic configurations. *Council for Research in Music Education Bulletin*, *22*, 4–12.

Engel, A., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top-down processing. *Nature Reviews Neuroscience*, *2*(10), 704–716.

Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Hogarth, L., Dickinson, A., Austin, A., Brown, C., & Duka, T. (2008). Attention and expectation in human predictive learning: The role of uncertainty. *Quarterly Journal of Experimental Psychology*, *61*(11), 1658–1668.

Huron, D. (1996). The melodic arch in western folksongs. *Computing in Musicology*, *10*, 3–23.

Huron, D. (2001). Tone and voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception*, *19*(1), 1–64.

Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.

Husserl, E. (1964). *The phenomenology of internal time-consciousness*. The Hague: Martinus Nijhoff. (Original work published 1905–1910)

Jackendoff, R. (1991). Musical parsing and musical affect. *Music Perception*, *9*(2), 199–230.

Jones, M.R. (1981). Music as a stimulus for psychological motion. Part I: Some determinants of expectancies. *Psychomusicology*, *1*, 34–51.

Jones, M.R. (1982). Music as a stimulus for psychological motion. Part II: An expectancy model. *Psychomusicology*, *2*, 1–13.

Jones, M.R. (1992). Attending to musical events. In M.R. Jones & S. Holleran (Eds.), *Cognitive bases of musical communication* (pp. 91–110). Washington, DC: American Psychological Association.

Koelsch, S., Maess, B., Grossmann, T., & Friederici, A. (2003). Electric brain responses reveal gender differences in music processing. *NeuroReport*, *14*(5), 709–713.

Krumhansl, C. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.

Kveraga, K., Ghuman, A., & Bar, M. (2007). Top-down predictions in the cognitive brain. *Brain and Cognition*, *65*(2), 145–168.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Margulis, E.H. (2003). Melodic expectation: A discussion and model. Unpublished doctoral dissertation, Columbia University.

Margulis, E.H. (2005). A model of melodic expectation. *Music Perception*, *22*(4), 663–713.

Margulis, E.H. (2007). Surprise and listening ahead: Analytic engagements with musical tendencies. *Music Theory Spectrum*, *29*(2), 197–217.

Meyer, L. (1956). *Emotion and meaning in music*. Chicago, IL: University of Chicago Press.

Meyer, L. (1967). *Music, the arts, and ideas*. Chicago, IL: University of Chicago Press.

Meyer, L. (2001). Music and emotion: Distinctions and uncertainties. In P. Juslin & J. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 341–60). Oxford: Oxford University Press.

Narmour, E. (1977). *Beyond Schenkerism: The need for alternatives in music analysis*. Chicago, IL: University of Chicago Press.

Narmour, E. (1990). *The analysis and cognition of basic melodic structures.* Chicago, IL: University of Chicago Press.
Narmour, E. (1992). *The analysis and cognition of melodic complexity.* Chicago, IL: University of Chicago Press.
Narmour, E. (1996). Analyzing form and measuring perceptual content in Mozart's Sonata K.282: A new theory of parametric analogues. *Music Perception*, *13*(3), 265–318.
Ockelford, A. (2005). *Repetition in music: Theoretical and metatheoretical perspectives.* London: Ashgate.
Ockelford, A. (2006). Implication and expectation in music: A zygonic model. *Psychology of Music*, *34*(1), 81–142.
Ockelford, A. (2008). Review of *Sweet anticipation: Music and the psychology of expectation*, by D. Huron. *Psychology of Music*, *36*(3), 367–382.
Ockelford, A. (2009). Similarity relations between groups of notes: Music-theoretical and music-psychological perspectives. *Musicae Scientiae*, Discussion Forum 4B, 47–93.
Olson, J., Roese, N., & Zanna, M. (1996). Expectancies. In E. Higgins &A. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 211–238). New York: Guilford Press.
Ortmann, O. (1926). On the melodic relativity of tones. *Psychological Monographs*, *35*(1), Whole No. 162.
Otten, M., Nieuwland, M., & van Berkum, J. (2007). Great expectations: Specific lexical anticipation influences spoken language processing. *BMC Neuroscience*, *8*(89), 1–9.
Panskepp, J. (1995). The emotional sources of "chills" induced by music. *Music Perception*, *13*(2), 171–208.
Povel, D.-J. (1996). Exploring the elementary harmonic forces in the tonal system. *Psychological Research*, *58*(4), 274–283.
Schellenberg, E. (1996). Expectancy in melody: Tests of the implication–realization model. *Cognition*, *58*(1), 75–125.
Schellenberg, E. (1997). Simplifying the implication–realization model of melodic expectancy. *Music Perception*, *14*(3), 295–318.
Schmuckler, M. (1989). Expectation in music: Investigation of melodic and harmonic processes. *Music Perception*, *7*(2), 109–150.
Sergeant, D., & Welch, G. (2009). Gender differences in long-term average spectra of children's singing voices. *Journal of Voice*, *23*(2), 319–336.
Steinbeis, N., Koelsch, S., & Sloboda, J. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience*, *18*(8), 1380–1393.
Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, *13*(9), 403–409.
Temperley, D. (1995). Motivic perception and modularity. *Music Perception*, *13*(2), 141–169.
Thompson, W., & Stainton, M. (1996). Expectancy in Bohemian folk song melodies: Evaluation of implicative principles for implicative and closural intervals. *Music Perception*, *15*(3), 231–252.
von Hippel, P., & Huron, D. (2000). Why do skips precede reversals? The effect of tessitura on melodic structure. *Music Perception*, *18*(1), 59–85.
Whittlesea, B., Masson, M., & Hughes, A. (2005). False memory following rapidly presented lists: The element of surprise. *Psychological Research*, *69*(5–6), 420–430.

Author biographies

Michael Thorpe is a doctoral student in the Department of Psychology at Roehampton University, London. His current work focuses on music cognition and perception in the domain of pitch.

Adam Ockelford is Co-Director of the Applied Music Research Centre at Roehampton University, London. His research interests are in music psychology, education, theory and aesthetics, particularly special educational needs and the development of exceptional abilities, the cognition of musical structure, and the construction of musical meaning. Forthcoming books include *Applied Musicology: Using Zygonic Theory to Inform Music Psychology, Education and Therapy Research (OUP) and Music, Language and Autism* (Jessica Kingsley).

Aleksandar Aksentijevic is a senior lecturer in Psychology at Roehampton University, London. He convenes a Masters programme in Applied Music Psychology. His research interests include auditory perception, psychology of music, dynamics of perceptual organization, and complexity.